Nucleic Acids Research, 2025, **53**, gkaf109
https://doi.org/10.1093/nar/gkaf109
Genome integrity, repair and replication

Transcription-replication conflicts drive R-loop-dependent nucleosome eviction and require DOT1L activity for transcription recovery

Marcel Werner[1,†], **Manuel Trauner**[1,†], **Tamas Schauer**[1], **Henning Ummethum**[1], **Elizabeth Márquez-Gómez**[1], **Maxime Lalonde**[1], **Clare S.K. Lee**[1], **Ioannis Tsirkas**[1], **Atiqa Sajid**[1], **Augusto C. Murriello**[1], **Gernot Längst**[2], **Stephan Hamperl**[1,*]

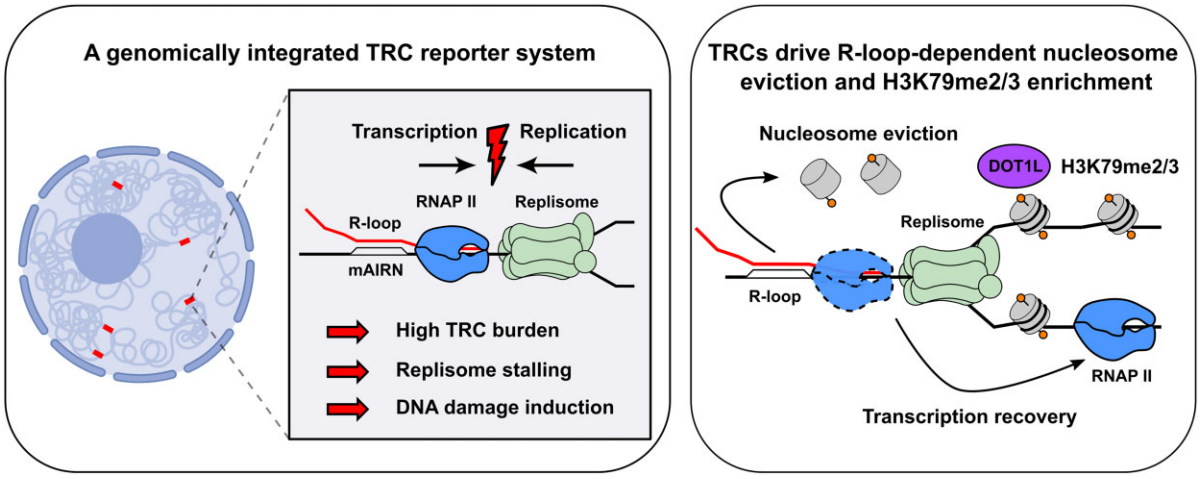
[1]Chromosome Dynamics and Genome Stability, Institute of Epigenetics and Stem Cells, Helmholtz Munich, Feodor-Lynen-Strasse 21, 81377 München, Germany
[2]Biochemistry Center Regensburg, University of Regensburg, Universitätsstr. 31, 93053 Regensburg, Germany

[*]To whom correspondence should be addressed. Email: stephan.hamperl@helmholtz-munich.de
[†]The first two authors should be regarded as Joint First Authors.

Abstract

Progressing transcription and replication machineries profoundly impact their underlying chromatin template. Consequently, transcription-replication conflict (TRC) sites are vulnerable to chromatin and epigenome alterations, provoking genome instability. Here, we engineered an inducible TRC reporter system using a genome-integrated R-loop-prone sequence and characterized the dynamic changes of the local chromatin structure inflicted by TRCs, leading to reduced nucleosome occupancy and replication fork blockage. Strikingly, inducing a small number of TRCs on the genome results in a measurable global replication stress response. Furthermore, we find a TRC-dependent increase in H3K79 methylation specifically at the R-loop forming TRC site. Accordingly, inhibition of the H3K79 methyltransferase DOT1L leads to reduced transcriptional output and an exacerbated DNA damage response, suggesting that deposition of this mark is required for effective transcription recovery and resolution of TRCs. Our work shows the molecular dynamics and reveals a specific epigenetic modifier bookmarking TRC sites, relevant to cancer and other diseases.

Graphical abstract



Introduction

DNA replication is an essential cellular process ensuring the correct duplication of the genome in the S-phase of the cell cycle. While replicating the genome, DNA replication forks face numerous exogenous and endogenous obstacles including chemically or UV-damaged DNA, DNA-bound proteins, and DNA secondary structures [1]. Failure to overcome these impediments can quickly lead to fork stalling, under-replicated

Received: July 3, 2024. **Revised:** January 31, 2025. **Editorial Decision:** February 3, 2025. **Accepted:** February 9, 2025

DNA, genomic instability, and mutations, hallmarks of cancer and aging [2, 3]. One particularly frequent endogenous obstacle is the presence of transcribing RNA polymerases. The lack of coordination between transcription and replication can lead to transcription-replication conflicts (TRCs), a hazardous genomic event connected to fork stalling, DNA breakage, and mutations [4–6]. TRCs can occur in two orientations: head-on (HO) and co-directional (CD). Conflict orientation determines distinct types of DNA damage responses [7]. Crucially, HO TRCs have been demonstrated to elicit a more profound impact on genome stability in bacteria [8, 9], yeast [10, 11], and mammalian cells [7].

HO TRCs are closely linked to the formation of R-loops [7, 9]. R-loops are three-stranded nucleic acid structures in which the template strand anneals with the nascently transcribed RNA, thereby forming an RNA:DNA hybrid and an exposed single-stranded non-template DNA strand. R-loops are prevalent structures in eukaryotic genomes and have physiological roles such as facilitating immunoglobulin class switch recombination in B-cells, telomere maintenance, chromosome segregation, and gene regulation. Nevertheless, they can also be pathological for example by inducing fork stalling and genome instability [4, 12–14]. Evidence from episomal constructs in human cells shows that R-loops exacerbate physical DNA breaks and DNA damage responses at TRC sites [7]. While R-loops are significant impediments to replication fork progression, the precise mechanism of how they obstruct fork progression at TRC sites is not fully understood. In HO orientation, the RNA:DNA hybrid part of an R-loop is on the opposite strand from the incoming replicative helicase and does not block replication unless the displaced single-stranded DNA (ssDNA) strand can form additional fork-blocking secondary structures like G-quartets [15]. Alternatively, an R-loop in combination with its associated RNA polymerase complex could act as a direct roadblock to replication [16–18]. Another possibility is that RNA:DNA hybrids could interfere with fork restart mechanisms behind stalled forks [19, 20]. A recent electron microscopy study and evidence from yeast support the existence of RNA:DNA hybrids behind the replication fork that interfere with the generation of replication protein A (RPA)-coated ssDNA and subsequent post-replicative repair [21, 22].

To prevent the harmful effects of TRCs, cells must rapidly detect and resolve TRCs. Over the past decade, various TRC resolution pathways have been identified. Evidence, primarily from *Escherichia coli*, suggests that the replisome can bypass the RNA polymerase (RNAP) complex by repriming downstream of the obstruction [23, 24]. During repriming, either the newly synthesized RNA can serve as a primer for replication restart, as demonstrated for CD TRCs, or a new Okazaki fragment may be synthesized [23, 25]. Another approach involves the degradation and/or removal of RNAP from chromatin to enable the resumption of DNA replication. Pathways such as RNAPII ubiquitination for degradation [26, 27], as well as the transient removal of RNAPII followed by replisome passage and transcription restart, have been described [28, 29]. Alternatively, the replication fork can be transiently cleaved and subsequently re-ligated, allowing RNA polymerases to restart transcription and progress past the replication fork. This multistep process depends on several key proteins, including the helicases RECQ1 and RECQ5, the scaffold protein SLX4, the endonuclease MUS81/EME1, RAD52, DNA ligase IV, the DNA polymerase δ subunit POLD3, and the transcription elongation factor ELL for transcriptional recovery [16]. Finally, in many of these pathways, the replisome may undergo an intermediate remodeling process called fork reversal. This mechanism stabilizes and protects the replication fork [21], providing the cell with additional time to resolve conflict.

Transcription and replication are both nuclear machineries that require an extensive amount of chromatin remodeling while traveling on the genome. Consequently, collisions between both protein complexes are likely to disrupt local chromatin organization [5], potentially leading to genetic and epigenetic instability. Indeed, several studies have highlighted specific histone modifications and chromatin proteins as crucial mitigators of TRCs and R-loop resolution. Initial studies in yeast have uncovered histone H3 serine 10 phosphorylation (H3S10P) as a chromatin mark required for R-loop driven chromatin compaction and genomic instability signaling [30]. Consecutive work established that mammalian H3S10P accumulates in large upstream (∼1 Mb) domains around R-loop prone HO-TRCs [31]. In addition, H3 lysine 4 (H3K4) methylation was characterized as a mitigator of TRCs by acting as a transcription-deposited "speed bump" to slow down replication forks, thereby preventing TRC occurrence upon replication stress [32]. Independently, a dynamic and replication stress-dependent switch of H2AK119 crotonylation to H2AK119 ubiquitylation was shown to act as a mechanism to release RNA polymerase II (RNAPII) and suppress transcription in the vicinity of stalled replication forks. As a result, this reduces the occurrence of TRCs and associated R-loops and DNA double-strand breaks (DSBs) [33].

Besides histone modifications, several other chromatin proteins have been implicated in TRC resolution. In mammalian cells, the BRG1 subunit of the SWI/SNF chromatin remodeling complex can resolve R-loop induced TRCs through a likely cooperative action with the Fanconi Anemia (FA) pathway [34]. Interestingly, the integrator complex, primarily associated with transcription termination, was recently demonstrated to attenuate CD TRCs and associated genomic instability by removing stalled RNAPII to maintain replication fork progression [35]. Intriguingly, integrator subunits interact with the MCM2-7-helicase independent of transcription, providing evidence for alternative functions of chromatin factors in transcription-replication coordination. Finally, MYCN-driven RNA exosome recruitment has been shown to be crucial to prevent TRCs in MYCN-dependent neuroblastoma [36].

Although a recent correlative study of genome-wide datasets defined a first network of chromatin factors such as SMARCA5, INO80, and MTA2 enriched at potential R-loop prone genomic TRC sites [31], a direct mechanistic and comprehensive understanding of TRC-induced changes to the local chromatin structure is still missing. This gap in our knowledge is largely due to numerous methodological challenges in studying TRCs, including limited approaches for genome-wide TRC mapping [37], and the lack of appropriate model systems to study TRCs in the context of complex mammalian genomes [5]. Controlled approaches in bacteria [9] and yeast [11, 38] allow for precise manipulation but frequently lack applicability to more complex mammalian genomes. In previous work, we have established a plasmid-based TRC reporter system that has yielded crucial insight into HO and CD conflicts and their association with R-loops and DNA damage [7]. Despite its previous utility in studying TRCs, the artificial nature

of the plasmid system and limited resemblance to endogenous chromatin composition impair its capability to address TRC-dependent chromatin biology.

In the present study, we leverage the previously used R-loop forming portion of the mouse AIRN (mAIRN) gene [39] to create a genome-integrated, inducible R-loop-based human TRC reporter cell line. After demonstrating that our model can successfully generate R-loops and TRCs, we show that R-loop-driven TRCs lead to replication impairment and a disruption of nucleosome abundance in the vicinity of TRC sites. In addition to this locally confined chromatin perturbation, we found that the induction of only a few TRCs at the integrated reporter sites can provoke a global DNA damage response and Ataxia telangiectasia and Rad3 related (ATR)-signaling-dependent proliferation defects. We profiled the transcription and replication-dependent dynamics of multiple chromatin marks and found H3 lysine 79 (H3K79) di- and trimethylation particularly enriched at the TRC reporter site upon induction. Importantly, we demonstrate a similar enrichment of this mark at TRC-prone and R-loop forming regions in the native genomic context, potentially providing a new marker for endogenous R-loop associated TRC detection in human genomes. Lastly, we show that inhibition of the H3K79 methyltransferase DOT1L prevents methylation of H3K79 at TRC sites, hindering the productive resumption of RNAPII transcription, ultimately leading to persistent fork stalling and exacerbated DNA damage. Altogether, these data indicate that replication-impeding R-loops and/or RNAPII complexes have a direct impact on the epigenetic landscape at TRC sites and establish a link between TRC induction, chromatin changes, deregulated replication, and DNA damage signaling, which may underlie the molecular basis of how TRCs impact the genetic and epigenetic stability of our genome.

Materials and methods

Cell culture

HEK293 Tet-ON and U-2 OS Tet-ON cells (Supplementary Table S2) were grown in Dulbecco's Modified Eagle Medium (DMEM) (GIBCO) containing 10% fetal bovine serum (FBS), 2 mM L-glutamine, and penicillin/streptomycin, in a 5% CO_2 environment at 37°C.

Antibodies and Reagents

A detailed list of all antibodies (Supplementary Table S3), bacterial strains (Supplementary Table S4), chemicals and recombinant proteins (Supplementary Table S5), critical commercial assays (Supplementary Table S6), oligonucleotides (Supplementary Table S7), and plasmids (Supplementary Table S8) used in this study can be found in the Supplementary Information.

Genomic integration of reporter sequence

For the construction of the mAIRN sequence containing Sleeping Beauty vector, the plasmid K192 pSBtet_DNMT3A_P2A_dsRed2 was amplified by PCR using primers MT073_gib_bb_fwd and MT073_gib_bb_rev to create a linearized vector of the Sleeping Beauty backbone with added overlap sequences for Gibson Assembly. Similarly, a fragment containing the mAIRN reporter sequence including its promoter and the SV40 poly-A signal was amplified from plasmid K069_pSH36_1xLEXA using primers MT073_gib_ins_fwd and MT073_gib_ins_rev, thereby also creating overlap sequences for Gibson Assembly. The mAIRN fragment and the Sleeping Beauty vector backbone were joined by Gibson Assembly using the Gibson Assembly Cloning Kit (NEB, REF: E5510S) according to the manufacturer's instructions. An identical strategy was employed for the Enhanced Cyan Fluorescent Protein (ECFP) sequence originating from plasmid K031_pSH26_1xLEXA.

Approximately 200 000 U-2 OS cells per well were seeded in a 6-well plate. A mix of 200 ng of the mAIRN reporter sequence-containing sleeping beauty shuttle plasmid (K275) and 1800 ng of the transposase expression plasmid (K191) [40] was prepared in 100 µL of OptiMEM, while 3.5 µL of Lipofectamine 2000 was diluted in 100 µL of OptiMEM. Both solutions were incubated for 5 min at room temperature, then mixed and incubated for another 20 min at room temperature. The cell culture medium was then replaced with 1.8 mL of fresh DMEM, and the mix was gently added to the well in a spiral pattern. The cells were grown for 24 h, then the medium was changed to 2 mL of fresh DMEM containing 1 µg/mL Puromycin. The polyclonal cell population was diluted to a concentration of 0.5 cells/100 µL per well on multiple 96-well plates. Hence, monoclonal cell lines were derived, growing from a single cell under 1 µg/mL Puromycin selection for further characterization.

Plasmid and siRNA transfections

An appropriate number of cells were seeded into the respective 96-well, 6-well, or 15 cm plate to reach 40–60% confluency at the day of transfection. Plasmid DNA and Lipofectamine 2000 were diluted in OptiMEM according to the manufacturer's recommendations. Both solutions were incubated for 5 min at room temperature, then mixed and incubated for another 20 min at room temperature. The cell culture medium was replaced with fresh DMEM, and the mix was added dropwise to the cells. For transfections of siRNA, 5 nM siRNA and Lipofectamine RNAiMAX were diluted in OptiMEM according to the manufacturer's recommendations. Both solutions were incubated for 5 min at room temperature, then mixed by vortexing, incubated for another 20 min at room temperature, and then added to the cells as described above.

MNase assay

Cells were grown in 6-well plates to 70–90% confluency (one well per sample/condition). Wells were washed with 2 mL of 1× PBS and trypsinized with 200 µL of 0.25 % trypsin. Cells were incubated at 37°C for 5 min and then resuspended in 0.5 mL of 1× PBS. The cell suspension was transferred to microcentrifuge tubes. Wells were washed with an additional 0.5 mL of 1x PBS, and the wash solution was combined with the cell suspension. The combined suspension was centrifuged at 500 g for 5 min. The cell pellet was resuspended in 1 mL of MNase lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.5% NP-40, 0.15 mM spermine, 0.5 mM spermidine) supplemented with 0.1mM phenylmethylsulfonyl fluoride (PMSF). Following a 5 min incubation on ice, nuclei were pelleted by centrifugation at 500 g for 5 min and washed with fresh MNase lysis buffer containing 0.1mM PMSF. The nuclei were again centrifuged at 500 g for 5 min and resuspended in 600 µL of MNase digestion buffer (10 mM Tris-HCl, pH 7.4, 15 mM NaCl, 60 mM KCl, 0.15 mM spermine, 0.5 mM spermidine, 1 mM $CaCl_2$). Aliquots of 100 µL of the nuclei prepa-

ration were transferred to separate microcentrifuge tubes containing diluted MNase (0, 2.5, 25, 100, and 250 gel units). Samples were incubated for 5 min at 30°C. Subsequently, 100 μL of IRN (50 mM TRIS-HCl pH 8, 20 mM EDTA, 500 mM NaCl) buffer, 7.5 μL of Proteinase K (10 mg/ml), and 20 μL of 10% sodium dodecyl sulphate (SDS) were added to each sample, followed by incubation overnight at 37°C or for at least 2 h. Phenol/chloroform/isoamylalcohol (200 μL) was added to each sample. The mixture was vortexed to achieve a homogenous suspension and then centrifuged for 5–10 min at 13 000 g. The upper (aqueous) phase was transferred to new tubes. 1 μL RNase A (10mg/mL) was added to the tubes, and the samples were incubated for 2 h at 37°C. DNA extraction was repeated with 200 μL of chloroform. Three volumes of 100% ethanol (p.a.) and 1 μL of glycogen (10 mg/mL) were added to each sample. Following incubation at −20°C for at least 30 min, samples were centrifuged at 13 000 g for at least 30 min at 4°C. The supernatant was removed, and the pellet was washed with 200 μL of 70 % ethanol. Samples were centrifuged again for 10 min at 13 000 g and 4°C. The pellet was dried and resuspended in 25–30 μL of H_2O or TE buffer. Finally, 10–15 μL of each sample was analyzed on a 1.2 % agarose gel.

Southern blot

After electrophoresis, the DNA fragments from the gel were transferred onto a nylon membrane for subsequent hybridization and detection with a probe. For denaturation and neutralization, the gel was submerged and washed 2 × 15 min with a denaturing solution (0.5 M NaOH, 1.5 M NaCl) to denature the DNA, followed by 2 × 15 min neutralization incubation with transfer buffer (1 M NH_4OAc). During the gel incubations, a nylon membrane and Whatman paper were prepared for the transfer. The membrane was soaked in water, and a larger Whatman paper was soaked in transfer buffer (1 M NH_4OAc). For the Southern blot transfer assembly, a platform was placed in a tray containing transfer buffer. A wet Whatman paper, free of air bubbles, was placed on the platform. This paper served as a wick to draw the transfer buffer upwards through the gel. The gel was then carefully placed face-down on the wet Whatman paper. A pre-wet nylon membrane was positioned on top of the gel, ensuring no air bubbles were trapped. Three additional Whatman paper sheets were placed on the membrane, followed by a stack of paper towels and a light weight (approximately 0.5 kg). The entire assembly was left undisturbed overnight to allow complete transfer of the DNA fragments. Transfer times varied depending on the fragment size, with fragments up to 15 kb requiring ∼18 h. Following transfer, the membrane was UV crosslinked using the Stratalinker autocrosslinking function to covalently bind the DNA fragments, enhancing hybridization signals during subsequent detection steps. The membrane could then be dried and stored at room temperature for further analysis. DNA probes for hybridization were generated using the RadPrime DNA labeling system (Invitrogen, 18428–011) with incorporation of [α−32P] dATP (Hartmann Analytik) according to the instructions of the manufacturer. Membranes were prehybridized for 1 h at hybridization temperature (65°C) with 10–15 ml of hybridization buffer (2 × SSC, 0.5 M sodium phosphate buffer pH 7.2, 7 % SDS). Following prehybridization, the buffer was discarded and replaced with 15 ml of fresh prewarmed hybridization buffer. The probe, mixed with salmon sperm DNA (final concentration 100 μg/ml) and boiled for 5 min, was then added to the tube. Hybridization occurred overnight at 65°C with gentle rotation in a hybridization oven. Blots were washed once with 30 mL of 3 × SSC and 0.1 % SDS after hybridization. Stringency washes were performed at hybridization temperature with rotation, using three buffers in sequential order: 0.3 × SSC with 0.1 % SDS, 0.1 × SSC with 0.1% SDS, and lastly 0.1x SSC with 1.5% SDS. Each wash step was repeated twice for 15 min. Finally, blots were dried, stored at room temperature, exposed to phosphorimaging screens and read out using a Typhoon scanner.

In vitro reconstitution of RNA:DNA hybrids

A T7 promoter was amended to the mAIRN DNA sequence by PCR, using a primer including the extended T7 promoter sequence (TAA TAC GAC TCA CTA TAG GGA GA). The PCR fragment was purified by gel extraction and used for *in vitro* transcription of the mAIRN sequence. Preparative *in vitro* transcription reactions (1 mL reaction volume, 5 h at 37°C) were prepared with 100 μg of recombinant T7 Polymerase, 2 μg of PCR-template and 5 mM NTPs in the reaction buffer (30 mM Tris-HCl pH 8.0, 25 mM $MgCl_2$, 2 μg Spermidin, 10 mM DTT, 0.01 % Triton X-100). RNAs were separated on denaturing polyacrylamide gels (8 %) and the RNA was excised from the gel. Gel elution from the fragmented gel slice was performed in 400 mM NaCl solution for 3 h at 4°C. The supernatant was collected and purified by isopropanol precipitation, dissolved, and quantified. RNA:DNA hybrids were reconstituted by mixing the fluorescently labeled PCR fragment with the purified RNA at a 1:2 ratio. The reaction was denatured (5 min at 95°C) and slowly cooled down. Efficient assembly into RNA:DNA hybrids was verified by restriction enzyme and RNase H digestion.

In vitro nucleosome formation assay

In vitro nucleosome assembly by the slat dialysis method was performed as previously described [41]. Recombinant human histone octamers were mixed with the corresponding DNA and 100 ng/μL BSA in high salt buffer (10 mM Tris-HCl pH 7.6, 2 M NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, 0.05% Igepal CA-630) and placed into small dialysis chambers. Dialysis chambers were placed into a beaker with 300 mL of high salt buffer. Overnight, 3 L of low salt buffer (10 mM Tris-HCl pH 7.6, 50 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, 0.05% Igepal CA-630) were pumped into the beaker containing the high salt buffer, thereby slowly reducing the NaCl concentration. Reconstituted nucleosomes were analyzed by gel electrophoresis, using 6% Polyacrylamide (PAA) 0.4 × Tris-borate-EDTA (TBE) gels and visualized by ethidium bromide staining or fluorescence imaging.

Cell proliferation assay and crystal violet staining

U-2 OS cells were plated at 3000 cells per well in a 6-well plate. After 24 h, cells were treated with the respective drugs and placed in an Incucyte S3 Live-Cell Analysis System. Cell proliferation was analyzed for a total duration of seven days. Obtained confluency measurements were normalized to the confluency at the 0 h measurement time point.

For crystal violet staining, cells were washed with PBS and fixed with 4% paraformaldehyde in PBS for 15 min in the dark. Cells were stained with crystal violet solution for 20 min with gentle agitation. Stained cells were washed three times

with PBS for 5 min with gentle agitation. The plates were dried at room temperature prior to imaging with a commercial camera.

Cell cycle synchronization (double thymidine block)

U-2 OS cells were seeded at a density of 1 million cells per 10 cm dish and cultured under the previously described conditions. The following day, the culture medium was changed to fresh DMEM medium containing 2 mM of thymidine. After 18 h, the cells were washed with PBS and fresh DMEM medium without thymidine was added. Nine hours later, the medium was again replaced with fresh DMEM medium containing 2 mM of thymidine. The cell arrest could either be released by adding fresh DMEM medium or continued by switching to fresh DMEM medium with 2 mM of thymidine. At the desired time point and cell cycle stage, cells were harvested using trypsinization and centrifugation.

Western blot

Samples were analyzed using a NuPAGE 4 to 12 % Bis-Tris Mini Protein Gel (Thermo Fisher) with 15 wells. The gel was transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore) and then blocked with 3% BSA in PBS. The primary antibody was added to the membrane and incubated in 3% BSA at 4°C overnight. Antibodies against GAPDH (1:1000), ORC2 (1:1000), γH2A.X (1:200), H3K79me2 (1:2500) and H3K79me3 (1:2000) were used. The membrane was washed twice with PBS-T (PBS with 0.05% Tween-20) for 5 min each. The secondary antibody (1:10 000) was then added in 3% BSA in PBS and incubated for 1 h. After two additional washes with PBS-T for 5 min, the membrane was developed on a ChemiDoc Touch Imaging System (BioRad) using the SuperSignal Kit (Thermo Fisher).

Flow Cytometry

The cells were subjected to a 30 min pulse-labeling with 25 μM 5-Bromo-2′-deoxyuridine (BrdU) followed by a wash with PBS before being trypsinized and harvested. The samples were fixed with ice-cold 70% ethanol, then permeabilized with 0.25 % Triton X-100/PBS for 15 min on ice. To denature the DNA to single strands, the cells were incubated in 2 M HCl for 15 min at 25°C and washed with 100 mM sodium borate (pH 8.5). The cells were blocked in 1% BSA/PBS containing 0.1% Tween-20 for 15 min, then incubated with a primary BrdU antibody (1:100; BD bioscience) for 2 h. After three washes with PBS, the cells were incubated with AlexaFluoro-488 secondary antibody (1:500; Invitrogen) for 1 h and washed three times with PBS. Propidium iodide (0.01 mg/mL) and RNase A (0.02 mg/mL) were added and the cells were analyzed on a FACSMelody device to determine DNA content. The cell cycle profiles were determined using FlowJo software.

DNA copy number analysis by qPCR

About 2–4 × 10^5 cells per sample were harvested by trypsinization and centrifugation. The cell pellet was resuspended in 100 μL of TE buffer, then the following reagents for DNA extraction were added: 100 μL IRN buffer, 10 μL 10% SDS (final concentration: 0.5%), and 10 μL Proteinase K (10 mg/mL, final concentration: 45 μg/mL). The mixture was incubated at 37°C for 2 h or overnight. Then, 200 μL phenol:chloroform:isoamyl alcohol (25:24:1, v/v) was added, followed by vortexing and centrifugation at 16 100 g for 5 min. The upper aqueous phase (∼180 μL, containing DNA) was transferred into a new 1.5 mL tube. Next, 5 μL RNase A was added, and the mixture was incubated at 37°C for 1 h. Finally, 200 μL chloroform was added, followed by vortexing and centrifugation at 16 100 g for 5 min to separate the phases. The upper aqueous phase was transferred into a fresh 1.5 mL microcentrifuge tube, DNA was precipitated with ethanol, and the DNA pellet was resuspended in DNAse-free water. The DNA was then mixed with iTaq SYBR Green Supermix according to manufacturer's instructions and analyzed by quantitative PCR (qPCR) on a Roche Light Cycler 480 Instrument II.

ChIP-qPCR

About 5 million cells per sample were harvested by trypsinization. The resulting cell pellet was resuspended in 9.5 mL of 1% formaldehyde solution in PBS and agitated for 10 min. The formaldehyde was neutralized by adding 0.5 mL of 2.5 mM glycine (pH 7) and incubated for an additional 5 min. The cells were centrifuged at 453 g at 4°C for 5 min and resuspended in 900 μL of PBS with protease and phosphatase inhibitors (Thermo Scientific). The cell pellets were rapidly frozen in liquid nitrogen and stored at −80°C until use. Preparation of Pierce Protein G Magnetic Beads involved adding 50 μL of bead slurry to a new 1.5 mL tube containing PBS on ice. The beads were then washed twice with IP dilution buffer (20 mM TRIS-HCl pH 8, 2 mM EDTA, 150 mM NaCl, 1% Triton-X 100, 0.01% SDS) and stored at 4°C until further use. Cell pellets were resuspended in 0.5 mL of cell lysis buffer (10 mM TRIS-HCl pH 8, 10 mM NaCl, 0.2% Igepal CA-630) with protease and phosphatase inhibitors, and incubated at 4°C for 10 min. The nuclei were separated by spinning at 1700 g for 5 min at 4°C, then resuspended in 550 μL of nuclei lysis buffer (50 mM TRIS-HCl pH 8, 10 mM EDTA, 1% SDS) with phosphatase and protease inhibitors, and 550 μL of IP dilution buffer. The nuclei were sonicated to a target size of 300 bp using the Covaris E220 sonicator (10 min sonication with 140 Peak Incident Power, 5% Duty Factor, 200 Bursts per Cycle) and the sonicated DNA spun at 16 000 g for 5 min at 4°C. 10% of the volume was saved as input sample (IN), and the rest was transferred to a new tube containing prepared Pierce Protein G Magnetic Beads. The tubes were incubated overnight with the antibody of interest at 4°C with agitation. Antibodies against RNAPII (8WG16) (5 μg), Histone H3 (4 μg), FANCD2 (5 μg), RNAPII Ser2P (10 μg), H3K4me3 (2 μg), H2AK119ub (2.5 μg), H3K79me2 (7.5 μg), and H3K79me3 (2 μg) were used in Chromatin immunoprecipitation (ChIP) experiments. The next day, multiple washes were performed at 4°C with agitation: one wash with IP wash 1 buffer (20 mM Tris-HCl pH 8, 2 mM EDTA, 50 mM NaCl, 1% Triton-X100, 0.1% SDS), two washes with high-salt buffer (20 mM Tris-HCl pH 8, 2 mM EDTA, 500 mM NaCl, 1% Triton-X 100, 0.1% SDS), one wash with IP wash 2 buffer (10 mM Tris-HCl pH 8, 1 mM EDTA, 250 mM LiCl, 1% NP-40, 1% deoxycholic acid), and two washes with TE (10 mM TRIS-HCl pH 8, 2 mM EDTA). The beads were resuspended in 47 μL of EB (50 mM TRIS-HCl pH 8, 10 mM EDTA, 1% SDS) and 3 μL of Protease K (10 mg/mL) and incubated at 56°C for 20 min. The supernatant was collected, and the process was repeated, with the supernatants being com-

bined. Around 6 μL of Proteinase K (10 mg/mL) was added to the IN samples and 100 μL of IRN buffer was added to all samples, which were incubated at 65°C overnight. DNA fragments were purified by phenol/chloroform extraction followed by ethanol precipitation and qPCR was performed as described.

DNA:RNA immunoprecipitation

Cells were lysed in 1.6 mL TE buffer supplemented with 82 μL of 10% SDS and 10 μL of 10 mg/mL Proteinase K and incubated at 37°C overnight. DNA was isolated by phenol:chloroform:isoamyl alcohol (25.24:1, v/v) extraction and isopropanol precipitation. DNA was reconstituted in 130 μL TE buffer, transferred to AFA microTube with Snap Cap and sonicated for 4 min using a Covaris E220 sonicator (140 peak incident power, 10% duty factor, 200 bursts per cycle). Sonicated DNA was quantified on a NanoDrop 2000c spectrophotometer. For RNase H-treated samples, 4 μg of DNA were treated with RNase H overnight at 37°C. For immunoprecipitation, 4 μg of DNA were bound to 6 μg of S9.6 antibody in 1X binding buffer (10 mM Na_3PO_4 pH 7, 140 mM NaCl, 0.05% Triton X-100) overnight at 4°C. Protein A/G agarose beads were added for 2 h. Bound beads were washed three times in 1X binding buffer for 10 min at 4°C. Elution was performed in elution buffer (50 mM Tris-HCl pH 8, 10 mM EDTA, 0.5% SDS, Proteinase K) for 45 min at 55°C with agitation.

Whole genome sequencing and identification of genomic integration sites (seq)

Genomic DNA was isolated by phenol:chloroform:isoamyl alcohol (25.24:1, v/v) extraction and ethanol precipitation. Genomic DNA was sonicated on a Covaris E220 sonicator (140 peak incident power, 10% duty factor, 200 bursts per cycle) to obtain DNA fragments of 200–300 bp size. Fragmented DNA was subjected to DNA library preparation procedure (for details see section on library preparation and sequencing). For integration site identification, libraries were sequenced to a depth of 300 million reads per sample.

BrdU-seq

BrdU-seq was adapted from a previously published BrdU-seq protocol [42]. For labeling of nascent DNA, cells were pulsed with 50 μM BrdU for 30 min before harvesting. Cells were lysed with 100 μL TE buffer, 100 μL IRN buffer, 0.5 % SDS, and 10 μL Proteinase K (10 mg/mL) and incubated at 37°C with 500 rpm agitation for 90 min. Genomic DNA was isolated by phenol:chloroform:isoamyl alcohol (25.24:1, v/v) extraction and ethanol precipitation. Reconstituted DNA was quantified by Qubit dsDNA BR kit (Thermo Scientific) on a Qubit 4 Fluorometer (Thermo Scientific) according to manufacturer's instructions. 20 μg of genomic DNA in 130 μL water were transferred to AFA microTubes with Snap Cap and sonicated for 10 min using a Covaris E220 sonicator (140 Peak Incident Power, 10% Duty Factor, 200 Bursts per Cycle) to obtain DNA fragments of 200–300 bp size. For BrdU immunoprecipitation, 10 μL Dynabeads Protein G were washed three times with PBS-T 0.02% and then incubated with 2 μg anti-BrdU antibody at 4°C overnight with gentle agitation. Sonicated DNA was diluted to 200 μL total volume, denatured at 100°C for 10 min and chased on ice to obtain ssDNA. 15 μL of denatured DNA were set aside as an input control and mixed with 35 μL elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS). 170 μL of denatured DNA were mixed with 180 μL of 2X blocking solution (2% BSA, 2X PBS, 0.2 % Tween 20) and added to the previously conjugated beads for immunoprecipitation at 4°C overnight with agitation. BrdU-labeled DNA bound to beads was washed twice with 1 mL of lysis buffer 1 (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-Deoxycholate). Next, samples were washed twice with 1 mL of lysis buffer 2 (50 mM HEPES-KOH pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1 % Na-Deoxycholate), followed by two washes with 1 mL of wash buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl,1 mM EDTA, 0.5% Na-Deoxycholate, 0.5 % Igepal CA-630). All washes were performed for a duration of 5 min at 4°C with gentle agitation. Beads were resuspended in 1 mL of TE and centrifuged for 3 min, 2000 rpm at 4°C. BrdU labeled DNA was eluted using 100 μL elution buffer with 0.5 mg/μL Proteinase K and incubated at 65°C for 10 min at 400 rpm. Following a repetition of the previous elution step, eluates were pooled and incubated at 37°C for 1 h. Input samples were supplemented with 150 μL TE buffer 0.5 mg/ml Proteinase K and incubated at 37°C for 1 h. Both BrdU-labeled DNA and input DNA were purified by phenol:chloroform:isoamyl alcohol (25.24:1, v/v) extraction and ethanol precipitation and reconstituted in 30 μL of water. Successful BrdU enrichment was verified by qPCR analysis. Second strand synthesis for library preparation was performed with RadPrime labeling kit (Thermo Scientific) according to the manufacturer's instruction. Obtained dsDNA was purified using the GeneJET PCR Purification Kit (Thermo Scientific).

Library preparation and sequencing

DNA samples were quantified with Qubit dsDNA HS Assay Kit (Thermo Scientific). Library preparation for Whole Genome Sequencing, ChIP-seq and BrdU-seq samples was performed according to manufacturer's instructions using the NEBNext Ultra II DNA Library Kit with Purification Beads (E7103S, E7103L) and NEBNext Multiplex Oligos for Illumina (E6440S). Libraries were pooled and sequenced on an Illumina NovaSeq 6000 machine with 150 bp paired-end reads at the Helmholtz Munich Genomics Core Facility.

RT-qPCR

About 2–4 × 10^5 cells per sample were harvested by trypsinization and centrifugation. The cell pellet was resuspended in 1 mL Trizol reagent and incubated at RT for 5 min. Then, 200 μL chloroform was added, and the mixture was vortexed for 10 s before incubating again at RT for 2 min. After centrifugation at 13 500 g at 4°C for 10 min, the upper aqueous phase (∼550 μL, containing DNA/RNA) was transferred into a new 1.5 mL tube. Next, 500 μL chloroform was added, vortexed for 10 s, and centrifuged at 13 500 g at 4°C for 10 min. The upper aqueous phase (∼500 μL, containing DNA/RNA) was transferred into a new 1.5 mL tube. Then, 500 μL isopropanol was added, vortexed for 10 s, and incubated at RT for 10 min. After centrifugation at 12 000 g at 4°C for 10 min, the supernatant was removed, and 1 mL 75% ethanol was added. Finally, it was centrifuged at 7500 g at 4°C for 15 min, and the pellet was allowed to air dry for a few minutes, then rehydrated in RNase-free water. Subsequently, cDNA synthesis was performed using the RT-SuperScriptIII

kit, following the manufacturer's instructions (ThermoFisher; Cat#18 080 051). qPCR was performed as described above.

EdU, immuno-, and DAPI staining

All steps were performed at room temperature unless otherwise specified. Cells were seeded in a 96-well plate (Ibidi, Cat#89 626) and treated with 10 μM EdU for 20 min before pre-extraction. Pre-extraction was done with CSK buffer (100 mM NaCl, 300 mM Sucrose, 3 mM $MgCl_2$, 10 mM MOPS, 0.5 % Triton X-100) for 5 min, then cells were washed with PBS and fixed with 4% PFA in PBS for 15 min. Following fixation, cells were washed with PBS, permeabilized with 0.2% Triton-X in PBS for 4 min and washed with PBS again. EdU click-it reaction solution (100 mM Tris-HCl pH 8.5, 1 mM $CuSO_4$, 100 mM Ascorbic acid, 0.9 μg Alexa Fluor 488) was added, and cells were incubated in the dark for 30 min. Cells were then washed with PBS and incubated in 3% BSA in PBS for 30 min. Primary antibody was applied in 3% BSA in PBS and the cells were incubated at 4°C overnight. Antibodies against FANCD2 (1:1000), and RNAPII RPB1 (H5) (1:2000) were used for immunofluorescence (IF) staining. The next day, cells were washed with 3% BSA in PBS and the secondary antibody and 4′,6-diamidino-2-phenylindole (DAPI) (1:1000) were applied in 3% BSA in PBS. The plate was wrapped in tin foil to protect from light and incubated for 1 h. Cells were then washed with PBS and stored at 4°C until needed.

Proximity ligation assays

All steps were done at room temperature unless stated otherwise. Cells were seeded in 96-well plates, pre-extracted, and fixed as previously described. They were then incubated in 5% BSA in PBS for 45 min, followed by application of primary antibodies in the same solution. Antibodies against RNAPII RPB1 (H5) (1:2000), PCNA (1:2000), and DOT1L (1:1000) were used in PLA. Cells were kept at 4°C overnight. The next day, cells were washed twice with PBS. Subsequently, Duolink PLUS (Sigma Aldrich) and MINUS (Sigma Aldrich) probes diluted with Duolink Antibody Diluent (1:10) were applied for 1 h at 37°C. The cells were washed twice with Wash buffer A (Sigma Aldrich) and the ligation solution (1x Duolink ligation buffer, Ligase at 1:70 dilution) (Sigma Aldrich) was applied for 30 min at 37°C. After two additional washes with Wash buffer A, the amplification solution (1x Amplification buffer, Polymerase at 1:140 dilution) (Sigma Aldrich) was added and cells were incubated in the dark at 37°C for 100 min. The cells were then washed twice with Wash buffer B (Sigma Aldrich) and stained with DAPI in 3% BSA in PBS for 1 h. Finally, cells were washed twice with PBS and stored at 4°C until imaging.

Microscopy

Images were acquired on a Nikon T2 inverted microscope equipped with an Andor Dragonfly spinning disk, a 40X air objective, and an iXon Life 888 EMCCD camera. Per condition 67–81 positions were imaged for which a Z-stack of 7 images across 10 μm was acquired.

Image analysis

A custom ImageJ [43, 44] macro was used to identify DAPI stained nuclei and perform multi-channel pixel-based mean intensity measurement and foci detection. In brief, the z-stack in focus was identified based on maximum normalized variance [45] in the DAPI channel. Following this, nuclei have been identified using a custom trained StarDist neural network model [46, 47].

Identifying EdU positive cells

S-phase cells were identified based on their nuclear mean EdU intensity. A threshold was set based on measured mean intensity levels in cell nuclei and by comparison with measured background level. Cells above a nuclear mean intensity of the threshold were considered as EdU positive cells.

Identification of insertion sites

The human genome (version GRCh38 primary assembly) was downloaded from the Ensembl database, and the sequence of the reporter construct was concatenated to the genome fasta file. Paired-end sequencing reads were aligned to the combined genome by bwa (version 0.7.17) with the parameter -B 3 [48]. Mapped reads were filtered by samtools (version 1.16.1) with the parameter -q 12 [49]. Structural variants were called on the filtered reads and the combined genome file using tiddit (version 3.3.2) with the parameters –sv -p 1 -r 1 –min_contig 4500 [50]. Genomic positions detected by reads split between the reporter and another chromosome were extracted from the variant call file (vcf). To exclude genomic positions that mapped internally to the reporter sequence, only those positions were considered that are within 300 bp of the reporter ends. Adjacent genomic positions with a maximum distance of 200 bp were merged using bedtools merge (version 2.31.0) with the parameter -d 200 [51] and the average of the start and end position of the merged region was used as insertion site. The same processing steps were carried out for the whole genome sequencing, the BrdU-seq, the RNAPII and H3 ChIP-seq datasets. The final insertion sites were identified as those genomic positions that were present in at least two of the four datasets within a 200 bp window (Supplementary Table S1). The latter step was carried out using rtracklayer (version 1.54.0) and GenomicRanges (version 1.46.1) [52] R packages (version 4.1.2).

BrdU-seq analysis

The combined genome with the reporter sequence was created as described above. Paired-end reads were trimmed using trim_galore (version 0.6.10) [53] with the parameter –quality 28 and mapped to the genome using bowtie2 (version 2.5.1) with the parameters –end-to-end –very-sensitive –no-unal –no-mixed –no-discordant –dovetail -I 10 -X 700 [54]. Aligned reads were filtered by samtools (version 1.17) with the parameter -q 12 [49]. Duplicated reads were removed using picard MarkDuplicates (version 3.0.0) with the parameter -REMOVE_DUPLICATES TRUE [55]. Genome coverages were created using deeptools bamCoverage (version 3.5.2) [56] with parameters –blackListFileName hg38-blacklist.v2.bed –ignoreForNormalization MT –binSize 20 –smoothLength 60 –extendReads –maxFragmentLength 700 –normalizeUsing "CPM". Blacklist regions were obtained from the Boyle lab github repository [57]. Reads were counted in 5 kb consecutive genomic bins using bedtools makewindows (version 2.31.0) with the parameter -w 5000 and deeptools multiBamSummary with parameters BED-file –BED bins.bed –blackListFileName hg38-blacklist.v2.bed –smartLabels –extendReads –centerReads –samFlagInclude 64 –outRawCounts. Only main chromosomes and the reporter

sequence were included in the counting step. Differential analysis was performed by DESeq2 package (version 1.34.0) [58] in R (version 4.1.2). For calculating normalization (size) factors all genomic bins were used by the default DESeq method. Differential testing was carried out in two different ways either globally for all genomic bins or in the proximity of the insertion sites, i.e. only for those bins that are within 100 kb of the insertion sites and for the reporter sequence. For the global analysis genomic bins with more than five reads counted in at least two out of all BrdU-seq samples were used whereas for the insertion proximal analysis more than one read count was required. Replicates were considered as batch variable in the DESeq model. For further analyses, normalized counts were log2 transformed after adding a pseudo-count of one and batch correction was applied using the ComBat function from the sva package (version 3.42.0) [59]. Global Spearman´s correlation coefficients were calculated on log2 normalized counts of all genomic bins. For heatmap visualization, log2 normalized counts were centered by the mean across samples for each bin and were plotted by pheatmap (version 1.0.12).

ChIP-seq analysis

Genome and ChIP-seq (RNAPII Ser2P and H3) read processing steps were carried out as described for BrdU-seq. In addition, coverages were averaged by deeptools bigwigAverage –binSize 20 and the log2 ratio between treatment and control was calculated by bigwigCompare –binSize 100 –fixedStep. Insertion site centered matrices were created by deeptools computeMatrix with parameters –referencePoint "center" –binSize 100 –downstream 5000 –upstream 5000 and were visualized using deeptools plotHeatmap function. Differential analysis was performed as for BrdU-seq, except, the bin size was set to 1 kb and more than four read counts per bin were required in at least half of the samples for the given ChIP antibody.

Public data analysis

Public datasets were obtained from GEO with the accession numbers GSE134084 (DRIP-seq) [60], GSE29611 (H3K79me2 ChIP-seq) [61], and GSE110354 (H3K79me3 ChIP-seq) [62]. Genome and read processing steps were carried out as described for BrdU-seq. In addition, peaks were called on DNA:RNA immunoprecipitation (DRIP)-seq replicates using macs2 (version 2.2.9.1) [63] with parameters –gsize 3e9 –format BAMPE –nomodel –nolambda –broad –broad-cutoff 1e-7 –pvalue 1e-7. Peaks overlapping blacklist regions were filtered out using bedtools intersect with parameter -v. The intersect of the replicates were defined as final peak set. ChIP-seq reads were counted similarly to BrdU-seq in 5 kb consecutive genomic bins. Bins with more than one read in at least half the samples were used. No batch correction was applied on the log2 normalized counts. Bins were grouped by the overlap with the DRIP-seq peaks and log2 normalized counts were visualized as box plots.

For the analysis of ENCODE ChIP-Seq data in Fig. 5H and Supplementary S5C–E, bigwig files were downloaded from GenCode. Additional bigwig files for ChIP-seq data on H3K79me2, RNAPII, and H3K79me2 in MOLM13 cells (GSE185094 – Samples GSM5606327, GSM5606329, and GSM5606331), H3K79me3 in HeLa cells (GSE116310 – Samples GSM3227896 and GSM3227897), and H2AK119ub in MCF7 cells (GSE201262 – Samples GSM6056895 and GSM6056896) were downloaded from NCBI GEO. Histone modification patterns around HO and CD collision regions identified by [7] were investigated. DeepTools computeMatrix tool was employed to analyze enrichment within 12 kb upstream and downstream of origins, using a 100 bp bin size. To assess enrichment significance specific to HO versus CD regions, 2000 bootstrap replicates were performed, and signals were visualized with correspondent 95 % confidence intervals.

Quantification and statistical analyses

Error bars on Figs 1E and F, 2C, E, F, H, J, K, and L, 3B, 4A and B, 5C–F, 6B, D, F, and G, Supplementary Figs S1A and B, S2C, E, G–L, N, and O, S3A–C, S4A, B, D, E, F, and G, S5A, S6A, B, and G, S7B–D, F, and H–L indicate the standard deviation (SD) of the biological replicates. All statistical details of the experiments can be found in the figure legends. All statistical comparisons were done using GraphPad Prism 10.0.2.

Statistical analysis of the BrdU-Seq and ChIP-Seq sequencing results in Figs 3 and 4, Supplementary Figs S3 and S5 was performed as described in the 'Materials and methods' section (see BrdU-seq analysis and ChIP-seq analysis). Detailed information about software and algorithms used for sequencing data including their sources and weblinks can be found in Supplementary Table S9.

Results

RNA:DNA hybrids are resistant to chromatin assembly and form an open nucleosome-depleted chromatin structure

In order to dissect the interplay between TRCs, R-loops, and nucleosome occupancy, we first reconstituted the R-loop prone mAIRN DNA sequence with its complementary RNA to form an RNA:DNA hybrid *in vitro*. As expected, the resulting mAIRN RNA:DNA hybrid induced a topological change that reduces mobility in native agarose gel electrophoresis (Fig. 1A). This shift was sensitive to RNase H treatment, confirming successful RNA:DNA hybrid reconstitution (Fig. 1A). In order to test the ability of this *in vitro* reconstituted RNA:DNA hybrid to form nucleosomes, we subjected the mAIRN hybrid as well as mAIRN double-stranded DNA (dsDNA) to a competition assay to form nucleosomes with the four core histone proteins. In contrast to dsDNA, RNA:DNA hybrids were unable to form nucleosomes despite supplementation of increasing amounts of histones, even though they consist of the same sequence as the dsDNA (Fig. 1B). Even at high histone octamer concentrations that result in precipitation of DNA, no binding of the histone octamers to the hybrid molecules was observed (Fig. 1B). This striking result indicates that the intermediate RNA:DNA hybrid duplex conformation is not compatible with nucleosome formation, potentially due to the stronger rigidity of the RNA strand than the DNA strand [64].

To investigate the interplay of R-loops and chromatin *in vivo*, we first took advantage of a previously established episomal system in human cells [7]. In brief, the constructs contain the doxycycline (DOX)-inducible (Tet-ON), R-loop-forming mAIRN transcription unit and the unidirectional origin of replication of the Epstein Bar virus (oriP) allowing for either HO or CD TRC induction (Fig. 1C). Previous analysis of R-loop levels on the mAIRN HO/CD constructs by DRIP and qPCR showed that DNA replication in HO

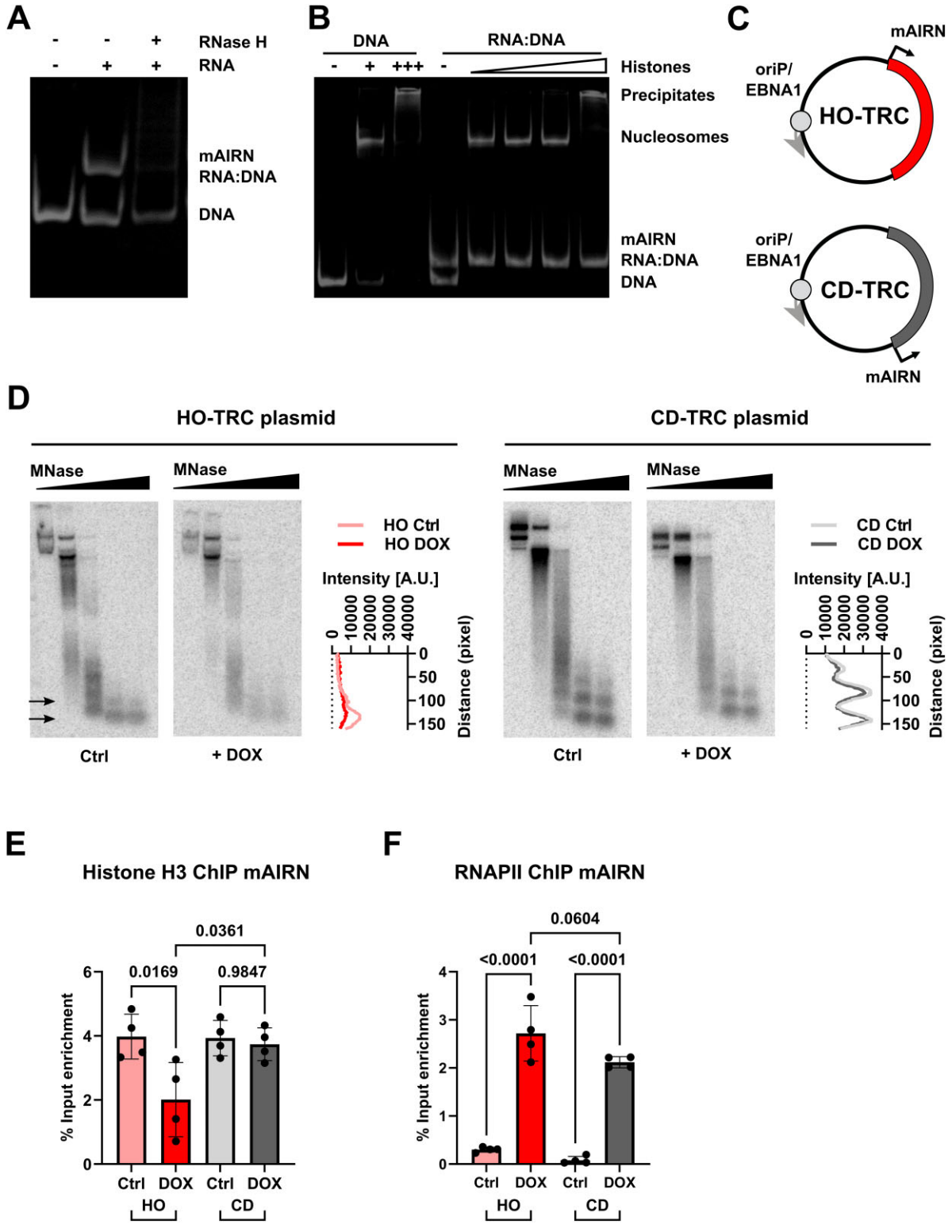


Figure 1. R–loop structures are incompatible with nucleosome formation *in vitro* and at an episomal HO-TRC reporter system *in vivo*. (**A**) Native polyacrylamide gel (6%) showing the reconstituted RNA:DNA hybrids (lane 2) compared to DNA alone (lane 1). Reconstituted RNA:DNA hybrids were treated with RNase H (5 U RNaseH, 30 min at 37°C) to indicate the specificity of the observed size shift (lane 3). (**B**) Native polyacrylamide gel (6 %) of the *in vitro* nucleosome assembly reactions on mAIRN dsDNA (lanes 1–3) versus mAIRN RNA:DNA hybrids (lanes 4–8) using increasing amounts of histone octamers, as indicated. The positions of the DNA, RNA:DNA hybrid, nucleosomes and over-assembled DNA-histone precipitates are indicated. (**C**) Schematic representation of the HO/CD TRC plasmid constructs. (**D**) Southern blot images of mAIRN HO/CD TRC plasmid constructs after treatment with 0 or 1 μg/ml DOX for 24 h. Samples were incubated with increasing concentration of MNase (0, 2.5, 25, 100, or 250 gel units of MNase). Arrows indicate nucleosome monomers and dimers. Quantification of Southern blot signal from nucleosome bands in arbitrary units (A.U.) for the 250 gel units MNase lane shown next to the blot. (**E**) ChIP-qPCR analysis for histone H3 at the mAIRN gene in HO and CD orientation treated with 0 or 1 μg/mL DOX for 24 h ($n = 4$). Ordinary one-way ANOVA with Tukey's multiple comparison test (**F**) ChIP-qPCR analysis for RNAPII at the mAIRN gene in HO and CD orientation treated with 0 or 1 μg/mL DOX for 24 h ($n = 4$). Ordinary one-way ANOVA with Tukey's multiple comparison test.

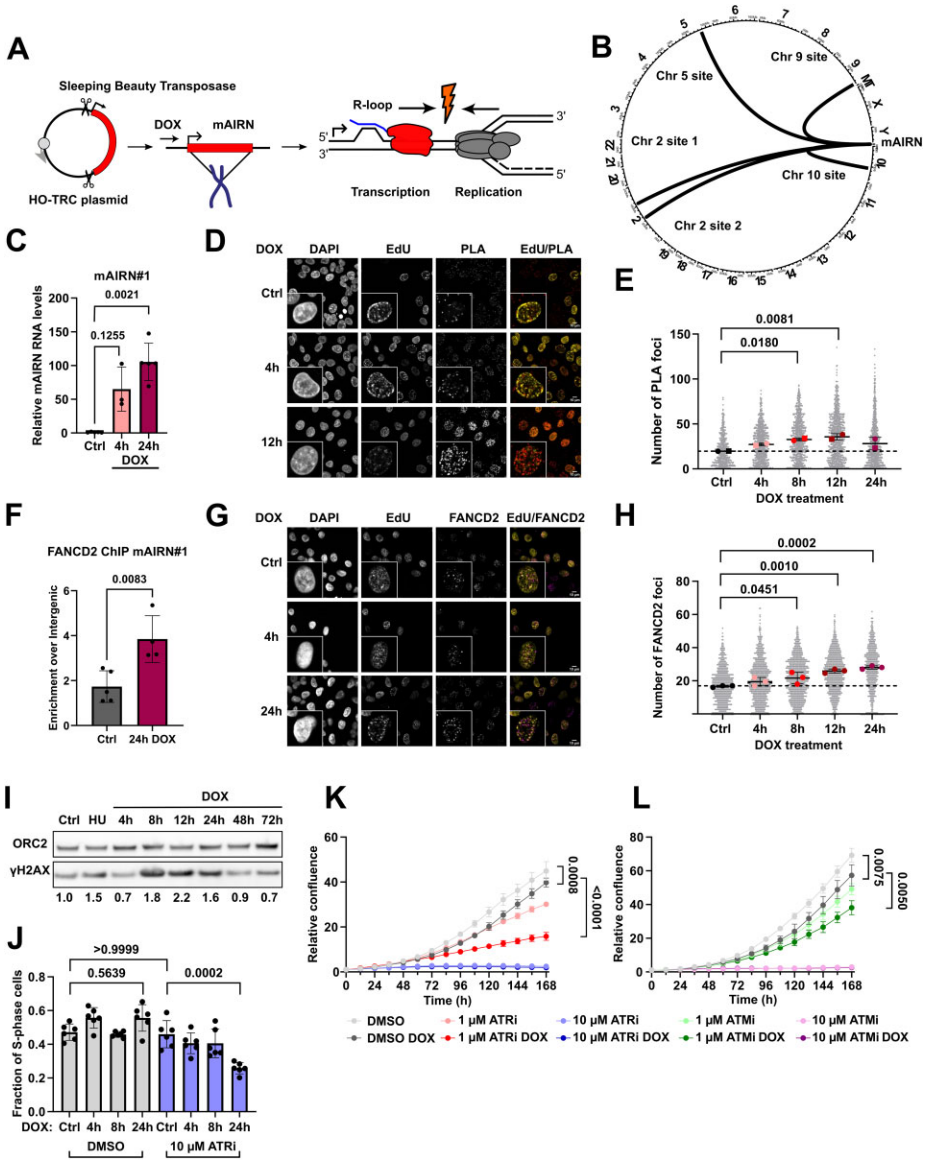


Figure 2. Chromosomal integration of an inducible R-loop forming gene increases cellular TRC levels and imposes a global replication stress response. (**A**) Diagram showing the generation of the chromosomal TRC reporter cell lines based on the R-loop forming mAIRN sequence of the episomal reporter plasmids. Genomic integration was achieved using Sleeping Beauty transposase integration. Inducible R-loop formation with DOX stalls RNAPII progression and strongly increases the occurrence of TRC events. (**B**) Circus plot showing the position of the identified integration sites of the TRC reporter construct in the monoclonal U-2 OS cell line clone#12. If not otherwise stated, clone#12 cells were used for all analyses throughout this study. Five integration sites on four chromosomes (Chr 2 site 1, Chr 2 site 2, Chr 5 site, Chr 9 site, and Chr 10 site) were found using whole genome sequencing followed by TIDDIT algorithm based structural variant calling. (**C**) RT-qPCR analysis of mAIRN RNA expression using primer pair mAIRN#1 in cells treated with 0 or 1 μg/mL DOX for 4 or 24 h ($n = 3$). Error bars indicate mean values with SDs. Welch ANOVA with Dunnett T3 comparison test (**D**) Representative images of TRC PLA assay with RNAPII Ser2P and PCNA antibodies (Ctrl, 4 and 12 h time points). EdU click-it staining was performed to label S-phase cells. Cells were treated with 0 or 1 μg/mL DOX for TRC induction. Scale bar 10 μm. (**E**) Quantification of TRC PLA foci number in S-phase cells from (D) as well as additional time points ($n = 2$, mean foci values per replicate as colored dots). Bars indicate mean values with SDs. Ordinary one-way ANOVA with Tukey's multiple comparison test. (**F**) ChIP-qPCR analysis showing FANCD2 levels in asynchronous cells at the TRC reporter sequence using the mAIRN#1 primer pair. Cells were treated with 0 or 1 μg/mL DOX for 24 h ($n = 4$). Error bars indicate SD. Unpaired t-test. (**G**) Representative images of FANCD2 IF staining (Ctrl, 4 and 24 h time points). Additional EdU click-it staining was performed to label S-phase cells. Cells were treated with 0 or 1 μg/mL DOX for TRC induction. Scale bar 10 μm. (**H**) Quantification of FANCD2 foci number in S-phase cells from (G) as well as additional time points ($n = 3$). Bars indicate mean values with SDs. Ordinary one-way ANOVA with Tukey's multiple comparison test (**I**) Representative Western blot analysis of DNA damage marker γH2AX using cell lysates from reporter cells treated with 0 or 1 μg/μL DOX for the indicated time points (0–72 h). 5mM Hydroxy urea (HU) was used as a positive control for DNA damage induction. ORC2 was used as loading control. Quantifications of γH2AX signal relative to the control condition shown below. (**J**) Quantification of fraction of S-phase cells based on EdU click-it staining in cells treated with 0 or 1 μg/mL DOX for 4, 8, or 24 h. 10 μM ATR inhibitor VE-821 (ATRi) or DMSO were additionally added as indicated. Control (Ctrl) cells were treated with ATRi or DMSO for 24h ($n = 6$). Bars indicate mean values with SD. Ordinary one-way ANOVA with Tukey's multiple comparison test. (**K**) Proliferation assay tracking the growth of TRC reporter cells upon treatment with 0 or 1 μg/mL DOX for a duration of 168 h in 12 h intervals using Incucyte S3 Live-Cell Analysis System. Cells were additionally challenged with 1 or 10 μM VE-821 (ATRi) or DMSO control treatment. Data points represent the mean of three replicates ($n = 3$) with error bars indicating SD. Area under the curve (AUC) measurements for each replicate were performed. Statistical analysis with Ordinary one-way ANOVA with Tukey's multiple comparison test was performed on the AUC measurements. (**L**) Proliferation assay identical to (K). Cells were additionally challenged with 1 or 10 μM KU-60019 (ATMi) or DMSO control treatment.

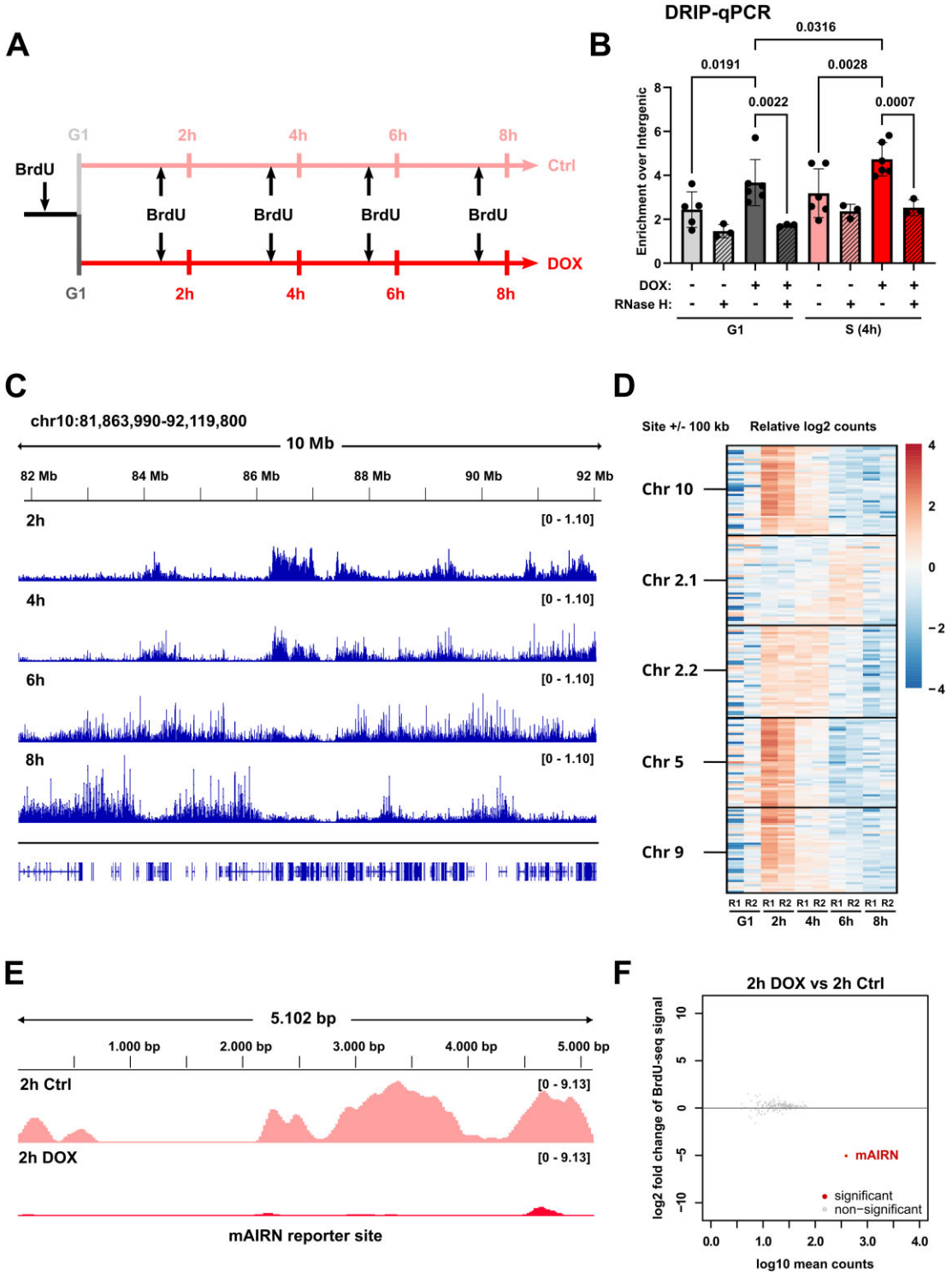


Figure 3. R-loop mediated TRCs slow or block DNA replication fork progression. (**A**) Treatment scheme of the BrdU-seq time course experiment. Cells were synchronized at the G1/S border using double thymidine block. Cells were released into S-phase for 2, 4, 6, or 8 h to allow S-phase progression. 30 minutes before harvesting, cells were pulsed with 25 μM BrdU for labeling of nascent DNA. Upon release, cells were treated with 0 or 1 μg/mL DOX. G1 cells were kept in thymidine conditions and treated with 0 or 1 μg/mL DOX for 8 h. (**B**) DRIP-qPCR analysis showing R-loop levels at the reporter sequence in G1 or 4 h released S-phase cells treated with 0 or 1 μg/mL DOX using the mAIRN#2 primer pair. For RNase H conditions, isolated genomic DNA from cells was treated with *E. coli* RNase H1 overnight to degrade R-loops ($n \geq 3$). Error bars indicate SD. Ordinary one-way ANOVA with Tukey's multiple comparison test. (**C**) Genome browser snapshot of a representative region on chromosome 10 indicating that BrdU-seq time course analysis tracks DNA replication progression and identifies early and late replicating domains. (**D**) Heatmap of BrdU-seq signal in ± 100 kb regions (5 kb bin size) around the integration sites in synchronized G1 cells and 2, 4, 6, and 8 h released S-phase cells. BrdU-seq signal is shown as log2 normalized read counts relative to the mean of all samples. Signal of both biological replicates is shown next to each other (R1 and R2). (**E**) Genome browser snapshot of BrdU-seq signal at the integrated mAIRN reporter construct at the 2 h S-phase time point in 0 or 1 μg/mL DOX treated cells, data shown from one of three biological replicates. (**F**) MA plot showing differential regulation of BrdU-seq signal at the 2 h S-phase release time point comparing DOX vs Ctrl conditions in the ± 100 kb regions around the integration sites shown in (D), in 5 kb bins. Significant and non-significant bins are highlighted accordingly.

orientation promotes R-loop formation, whereas R-loop levels are reduced in CD orientation [7]. Based on this difference in R-loop levels on the mAIRN HO and CD constructs, we tested how the specific enrichment of these secondary structures upon HO collision with the replication fork affects the chromatin of the episomal constructs. To this end, we performed limited Micrococcal nuclease (MNase) digestion and Southern blot analysis to determine nucleosome occupancy of the mAIRN HO and CD transcription units with and without DOX-induced transcriptional activation (Fig. 1D). In the absence of DOX, both mAIRN HO and CD constructs gave products corresponding to nucleosome monomers and dimers, suggesting that both plasmids are chromatinized in cells (Fig. 1D). Strikingly, DOX treatment induced specific loss of nucleosomes on the mAIRN HO construct, but not on the CD construct (Fig. 1D). This suggests that HO conflicts lead to nucleosome-depleted regions at the collision site. To further corroborate these results, we performed ChIP against canonical histone H3. Consistent with the MNase analysis, we observed specific reduction of H3 on the mAIRN sequence upon transcriptional activation of the HO construct (Fig. 1E). In agreement with the MNase assay, CD episomes retain histone H3, arguing for no changes in histone occupancy at the reporter sequence (Fig. 1E). As controls, we also performed ChIP against total RNAPII and observed similar recruitment of RNAPII on both HO and CD episomes upon DOX induction (Fig. 1F), confirming similar levels of RNAPII loading on both constructs. Importantly, an unrelated highly expressed genomic control locus ACTB and an intergenic region showed neither a DOX-dependent increase in RNAPII levels (Supplementary Fig. S1A), nor displayed any changes in H3 levels upon DOX treatment (Supplementary Fig. S1B). Taken together, our data supports the idea that mAIRN RNA:DNA hybrids and especially stable mAIRN RNA:DNA hybrids induced by HO TRCs can displace nucleosomes leading to a local chromatin disruption, thereby creating a potential susceptibility to DNA damage.

Chromosomal integration of an inducible R-loop forming gene increases cellular TRC levels and imposes a global replication stress response

Although the episomal reporter system shows clear TRC orientation-dependent disruption of chromatin, it is debatable to what extent the episomes are sufficiently chromatinized to reflect the dynamics of TRCs in the native chromatin context. Therefore, we next sought to develop a TRC reporter system stably integrated into the genome to address TRC-driven chromatin disruption. To this end, we cloned the previously used mAIRN sequence along with its DOX-inducible promoter into a Sleeping Beauty transposase vector [40, 65], allowing us to insert the sequence into the genome of U-2 osteosarcoma (U-2 OS) cells (Fig. 2A). Such random genomic integration with Sleeping Beauty also enables the simultaneous study of multiple insertions at different locations and under diverse chromatin contexts. For the construction of the model system, we exclusively focused on the R-loop forming mAIRN sequence, since the corresponding unidirectional origin oriP had been shown to be late replicating in mammalian cells [66, 67] and would likely not be able to initiate replication at a rate needed for efficient TRC induction. Despite this, we expected that upon induction of the Tet-ON promoter with DOX, the mAIRN-driven formation of stable R-loops in complex with stalled RNAPII would create a potent genomic obstacle, thereby strongly increasing the chance for collisions with approaching endogenous replication forks leading to TRCs (Fig. 2A).

We generated several monoclonal U-2 OS cell lines with insertions in different genomic locations. Six of the resulting clones were tested for copy number by qPCR (Supplementary Fig. S2A), from which several clones were subjected to whole genome sequencing to precisely identify the genomic locations of the mAIRN integration sites. As an example, one of the clones (clone #12) showed five integration sites across four different chromosomes (Fig. 2B), some of which we also verified by genotyping PCR and Sanger sequencing (Supplementary Fig. S2B). In contrast to the parental cell line, RT-qPCR analysis in clone #12 showed 50–100-fold transcriptional induction of the mAIRN reporter sequence after 4 and 24 h DOX treatment, respectively (Fig. 2C, Supplementary Fig. S2C). Unless stated otherwise, we chose the clone #12 stable cell line as the predominant working model for all subsequent experiments.

We first wondered if the R-loop-driven reporter sequence can indeed successfully induce TRCs. To this end, we conducted proximity ligation assay (PLA) as a proxy for TRC formation, using antibodies against actively elongating RNAPII (RNAPII Ser2P) and PCNA as a component of the replication fork [7]. Strikingly, reporter activation with DOX strongly increased the level of TRC-PLA foci in S-phase cells when compared to the basal state (Fig. 2D). Elevated TRC-PLA foci levels were visible as early as 4 h upon DOX induction and further increased significantly after 8 and 12 h of induction (Fig. 2E). Prolonged DOX exposure up to 24 h resulted in a moderate decrease in TRC-PLA foci, indicating potential compensatory mechanisms. Similar to this clone, another characterized cell line containing the identical mAIRN reporter at distinct genomic locations also demonstrated a higher TRC burden upon DOX induction (Supplementary Fig. S2D and E). In contrast, the parental U-2 OS cells without mAIRN insertions did not show a DOX-dependent increase of TRC-PLA foci (Supplementary Fig. S2F and G), supporting the notion that TRC induction is specific to the integrated mAIRN sequence independent of the particular clonal genomic context.

Previous studies have demonstrated that TRCs can cause genomic instability in mammalian cells [7,34,68,69]. Consequently, we wondered if TRCs generated on the mAIRN-driven genomic reporter can induce replication stress and DNA damage. FANCD2, a marker of stalled and damaged replication forks has previously been connected to R-loop and TRC induction [70–72]. Thus, we analyzed the enrichment of FANCD2 at mAIRN loci by ChIP-qPCR. Importantly, we benchmarked the specificity of our FANCD2 ChIP with aphidicolin (APH) treatment, which showed an increase of FANCD2 levels at the common fragile site (CFS) loci NRG3 and WWOX as previously reported (Supplementary Fig. S2H) [73]. Remarkably, DOX treatment for TRC induction significantly increased FANCD2 binding at the mAIRN reporter sequence (Fig. 2F), thereby providing direct evidence for a TRC locus-driven replication stress phenotype. DOX treatment did not affect FANCD2 levels at CFS genes (Supplementary Fig. S2H), thus indicating reporter site specificity. Furthermore, to test whether this phenotype is dependent on R-loop formation at mAIRN, we also created a control cell line that contains a DOX-inducible ECFP sequence instead of mAIRN in analogy to the original episomal TRC system [7]. Importantly, the selected monoclonal ECFP cell line (clone #2)

showed a comparable number of integration sites and DOX-dependent transcriptional activation to exclude confounding differences in transcription levels or integration site copy numbers (Supplementary Fig. S2I and J). In contrast to mAIRN, FANCD2 levels were not significantly affected in the ECFP gene upon DOX induction (Supplementary Fig. S2K), providing independent evidence that the observed phenotype is specific to the R-loop forming mAIRN TRC reporter loci.

To further explore this replication stress phenotype, we performed IF staining for FANCD2 combined with EdU-Click-it labeling for detection of S-phase cells. Strikingly, the number of chromatin-bound FANCD2 foci were globally increased upon R-loop driven TRC induction across the experimental time course with highest levels at 24 h induction time (Fig. 2G and H). Furthermore, FANCD2 levels increased in both S-phase and non-S-phase cells (Fig. 2H, Supplementary Fig. S2L). Although non S-phase FANCD2 foci could stem from post-replicative DNA repair or synthesis spots that could not be resolved within the same S-phase and persisted into G2/M or G1-phase as recently described for MiDAS [74] and G-MiDS [42], we focused on S-phase FANCD2 foci as these are more likely associated with DNA damage upon TRC induction. Importantly, global RNAPII Ser2P levels were unchanged over the same time course and independent of cell cycle state (Supplementary Fig. S2M and N). Western blot analysis confirmed a global induction of phosphorylated H2A.X (γH2AX) as replication stress marker compared to the nuclear loading control ORC2 (Fig. 2I), showing a global replication stress response upon TRC induction with DOX in our TRC reporter cell line. As γH2A.X is a substrate of the ATR DNA damage response kinase previously connected to TRCs [7, 36], we next challenged cells with a specific ATR inhibitor (ATRi). Strikingly, ATRi-treated cells showed sensitivity to TRC induction as indicated by a significant decrease in the fraction of S-phase cells after 24 h of TRC induction with DOX, whereas control cells without ATRi (DMSO) maintained their number of S-phase cells and overall EdU incorporation levels (Fig. 2J, Supplementary Fig. S2O). To corroborate these results, we analyzed the proliferation of our TRC reporter cell line after DOX induction in combination with ATRi (Fig. 2K, Supplementary Fig. S2P) or ATMi (Fig. 2L, Supplementary Fig. S2Q). DOX treatment alone showed a modest but significant defect in cell proliferation, which was strongly exacerbated in combination with 1 μM ATRi and to a lesser extent with ATMi treatment. Taken together, these results show that TRC induction causes ATR- and ATM-dependent proliferation defects. Moreover, our data indicate that cells can tolerate such increased TRC burden in principle, but only if the DNA damage response is fully functional.

R-loop mediated TRCs slow or block DNA replication fork progression

As elevated FANCD2 levels indicate an increase in replication fork stalling events, we wondered whether induced TRCs and R-loops directly impair DNA synthesis rates at the mAIRN reporter sites. To this end, we performed a double thymidine block to synchronize cells at the G1/S border and subsequently removed thymidine to release cells into S-phase for different time points from 0 to 8 h in the absence or presence of DOX (Fig. 3A). We confirmed successful cell synchronization by analyzing the incorporation of the nucleotide analog BrdU as a proxy for DNA synthesis by flow cytometry and found more than 98% of cells arrested in G1/S and successful S-phase release with more than 91% of cells released after 4 h of thymidine removal (Supplementary Fig. S3A).

Using this cell cycle synchronization strategy, we first analyzed R-loop levels at the mAIRN reporter sites using DRIP-qPCR in G1/S arrested and 4 h S-phase-released cells. R-loop levels were about 2-fold significantly increased following transcriptional induction of the reporter sequence with DOX in both G1 and S-phase cells and could be reduced to near baseline state by *in vitro* treatment with *E. coli* RNase H, confirming DRIP specificity (Fig. 3B). As an important control, unrelated R-loop-prone loci such as RPL13A were unaffected by reporter induction with DOX treatment (Supplementary Fig. S3B). Strikingly, S-phase cells exhibited higher R-loop levels when compared to G1-phase cells, despite highly similar levels of transcription induction (Supplementary Fig. S3C). This replication-dependent stabilization of R-loops might arise due to co-occurring TRCs in S-phase cells.

To characterize the effect of induced TRCs and R-loops on DNA replication across different S-phase stages, cells were subjected to a BrdU pulse 30 min before sample collection for each time point following the release from the double thymidine block. BrdU-labeled DNA was pulled down and subjected to BrdU-Seq. On the global level, the BrdU-Seq data showed very high correlation between individual replicates and clustered according to the different time points of S-phase release (Supplementary Fig. S3D). Using this time-resolved BrdU-Seq data, we then asked if R-loops and coinciding TRCs would impair replication fork progression at the reporter sequence and/or surrounding genomic regions close to the integration sites. As shown for an example 10 Mb region on Chr 10 (Fig. 3C), our experimental setup can reproducibly distinguish early, mid, and late replicating regions of the genome. Quantification of the BrdU-seq signal in a ± 100 kb window around the integration sites allowed us to define the replication timing for each of the five integration sites. Interestingly, Chr 5, Chr 9, and Chr 10 sites were all located in early replicating regions with the strongest BrdU-seq signal at 2 h into S-phase and little to no replication activity at later time points (Fig. 3D). Chr 2 site 1 exhibits a mid to late replication timing starting replication around 6 h after release. Chr 2 site 2 exhibits early to mid S-phase replication around 2–4 h after release. Control versus DOX-treated cells also highly correlated in all time points tested, suggesting that mAIRN transcription induction had no effect on the global DNA replication landscape of the cells (Supplementary Fig. S3D). Instead, we observed a strong reduction of BrdU-seq signal exclusively at the mAIRN reporter site (Fig. 3E and F, Supplementary Fig. S3E and F). Significant local disruption of BrdU incorporation was only observable at the 2 h time point, coinciding with the expected replication timing of four out of five integration sites. We additionally quantified the reduction of BrdU-seq signal using the same ±100 kb window around the sites as in Fig. 3D. Similarly, the ∼5 kb reporter sequence remained the only significantly changed genomic bin with reduced BrdU incorporation at the 2 h time point (Fig. 3F). Strikingly, there was no difference in BrdU incorporation at the subsequent 4, 6, and 8 h release time points (Supplementary Fig. S3G–I), although slightly elevated BrdU-seq signal at the induced reporter sequence at 4h might indicate compensatory DNA synthesis at mAIRN to complete the gap in DNA synthesis at 2 h. Collectively, our data suggest that induced TRC and mAIRN R-loop formation can directly cause strong local DNA replica-

tion fork impediments likely causative of the observed fork stalling, DNA damage, and proliferation phenotypes.

TRC induction disrupts local chromatin structure on integrated R-loop reporter sites

Since the reporter system-driven TRCs can cause replication impairments and DNA damage, we next wondered if this local interference would coincide with or cause alterations in the underlying chromatin template. In analogy to our episomal reporter system, we first sought to understand the dynamics of RNAPII and nucleosomes upon reporter gene activation. To this end, we performed ChIP-qPCR against RNAPII Ser2P and histone H3 in synchronized cells after 4 h S-phase release. As expected, RNAPII Ser2P levels were strongly increased upon DOX treatment confirming successful transcriptional activation of the mAIRN locus (Fig. 4A). Strikingly, transcription induction reduced nucleosome occupancy at the mAIRN sequence as indicated by a pronounced reduction of histone H3 enrichment at mAIRN loci (Fig. 4B). DOX treatment did not affect RNAPII Ser2P and histone H3 levels at unrelated control loci in mAIRN cells (Supplementary Fig. S4A and B). To further test whether the observed H3 loss is specific to R-loop formation, we attempted to rescue the nucleosome depletion and accumulation of FANCD2 by overexpression of human FLAG-tagged RNase H1 in mAIRN cells. Although Western blot analysis showed high levels of RNase H1 overexpression 24 h after transfection (Supplementary Fig. S4C), the results show that RNase H1 overexpression was unable to rescue both phenotypes at induced mAIRN loci (Supplementary Fig. S4D, Supplementary Fig. S4E).

We do not attribute this negative result to a lack of R-loop dependency of these effects, but rather to technical limitations that our RNase H1 overexpression construct was unable to efficiently degrade and/or remove the particularly strong and stable mAIRN R-loops (see 'Discussion'). However, to demonstrate R-loop specificity by a complementary approach, we took advantage of the control cell line with non-R-loop forming ECFP sequences (Supplementary Fig. S2I and J). In contrast to mAIRN cells, H3 levels were not significantly affected in the ECFP gene or unrelated control locus upon DOX induction (Supplementary Fig. S4F and G), providing independent evidence that the observed phenotype is specific to the R-loop forming mAIRN TRC reporter loci.

Subsequent ChIP-seq analysis of the average signal across the five integration sites in mAIRN cells revealed that the transcriptional activation and histone reduction are not confined to the mAIRN sequence but extend throughout the complete ∼5 kb reporter construct (Fig. 4C). As this analysis necessarily averages the RNAPII Ser2P and H3 signals across all five integration sites, we next wondered how activation of the individual reporter loci affects the neighboring chromatin directly upstream and downstream of the integration sites. Interestingly, RNAPII Ser2P recruitment upon DOX does not remain confined to the reporter for most integration sites. For example, Chr 2 site 1 shows elevated RNAPII Ser2P levels extending several thousand base pairs into the upstream intronic sequence of the MRPS9-AS2 gene (Fig. 4D). Consistently, RT-qPCR analysis showed ∼50-fold DOX-induced transcriptional activation of this intronic sequence (Supplementary Fig. S5A), whereas neighboring genes of Chr 2 site 2 and Chr 10 sites located further upstream or downstream of the readthrough transcription regions did not display gene expression changes (Supplementary Fig. S5A). As this DOX response was predominantly observed in one direction of the integration site, this is likely the result of readthrough transcription depending on the orientation of the integrated Tet-ON promoter driving mAIRN expression. Similar directional extension of RNAPII occupancy was observed for the other integration sites except for the site in Chr 9, where no RNAPII Ser2P induction was detected outside of the reporter sequence (Fig. 4E). DOX treatment did not induce global changes of RNAPII Ser2P levels except for genomic bins in the direct neighborhood of the integration sites (Supplementary Fig. S5B). Crucially, these adjacent regions (the actual ∼5kb integration sites are excluded from this visualization) exhibiting RNAPII readthrough transcription did not show a reduction of H3 (Fig. 4D and F), suggesting that the H3 loss phenotype is specific to the mAIRN reporter driving R-loop formation and TRC induction and cannot just be explained by transcription-induced changes in RNAPII occupancy. Again, DOX treatment did not alter H3 levels globally (Supplementary Fig. S5C–E). Collectively, our data show that R-loop-driven TRCs can locally reduce nucleosome occupancy, thereby disrupting the protective function of chromatin and increasing the susceptibility to DNA breaks and genome instability.

H3K79 methylation is a TRC-enriched chromatin modification at the R-loop reporter and genome-wide

Apart from the increased chromatin accessibility due to nucleosome eviction, we also investigated whether TRC induction has an impact on the local histone post-translational modification (PTM) state of the R-loop forming reporter gene chromatin. Several histone PTMs including H3K4 methylation, H3S10 phosphorylation, and H2AK119 ubiquitination have previously been linked to TRC biology in different cellular contexts and model systems [31–33, 75], providing us with a list of candidate histone PTMs that we profiled by ChIP-qPCR after TRC induction in our reporter cell line. Importantly, we included for each modification tested, a side-by-side H3 ChIP from the same cell lysate to account for the observed nucleosome loss upon transcription activation of the mAIRN reporter (Fig. 5A). As expected, H3 reduction was clearly observed in G1-arrested cells after 4 h of DOX induction along multiple primer pair locations of the reporter sequence and further decreased upon release for 4 h into S-phase in the presence of DOX (Fig. 5B, Supplementary Fig. S6A, TSS, mAIRN#1, mAIRN#2). This effect was not observed at the unrelated control NRXN2 gene (Fig. 5B, Supplementary Fig. S6B, NRXN2). H3K4 trimethylation (H3K4me3) marks transcription start sites with a 5′ to 3′ gradient of decreasing H3K4 methylation into the gene body and has been recently shown to mitigate TRCs by decelerating ongoing replication [32]. Consistently, we observed an S-phase DOX-dependent increase in H3K4me3 levels at two primer locations in the mAIRN reporter gene body (Fig. 5C, mAIRN#1 and mAIRN#2), but not upstream at the TSS or the control NRXN2 gene (Fig. 5C, TSS and NRXN2). Furthermore, it was recently shown that dynamic switching of crotonylation to ubiquitination of H2A at lysine 119 (H2AK119ub) leads to the release of RNAPII and repression of TRCs [33]. In agreement, H2AK119ub levels increased upon TRC induction in S-phase cells at the mAIRN reporter gene, most notable at

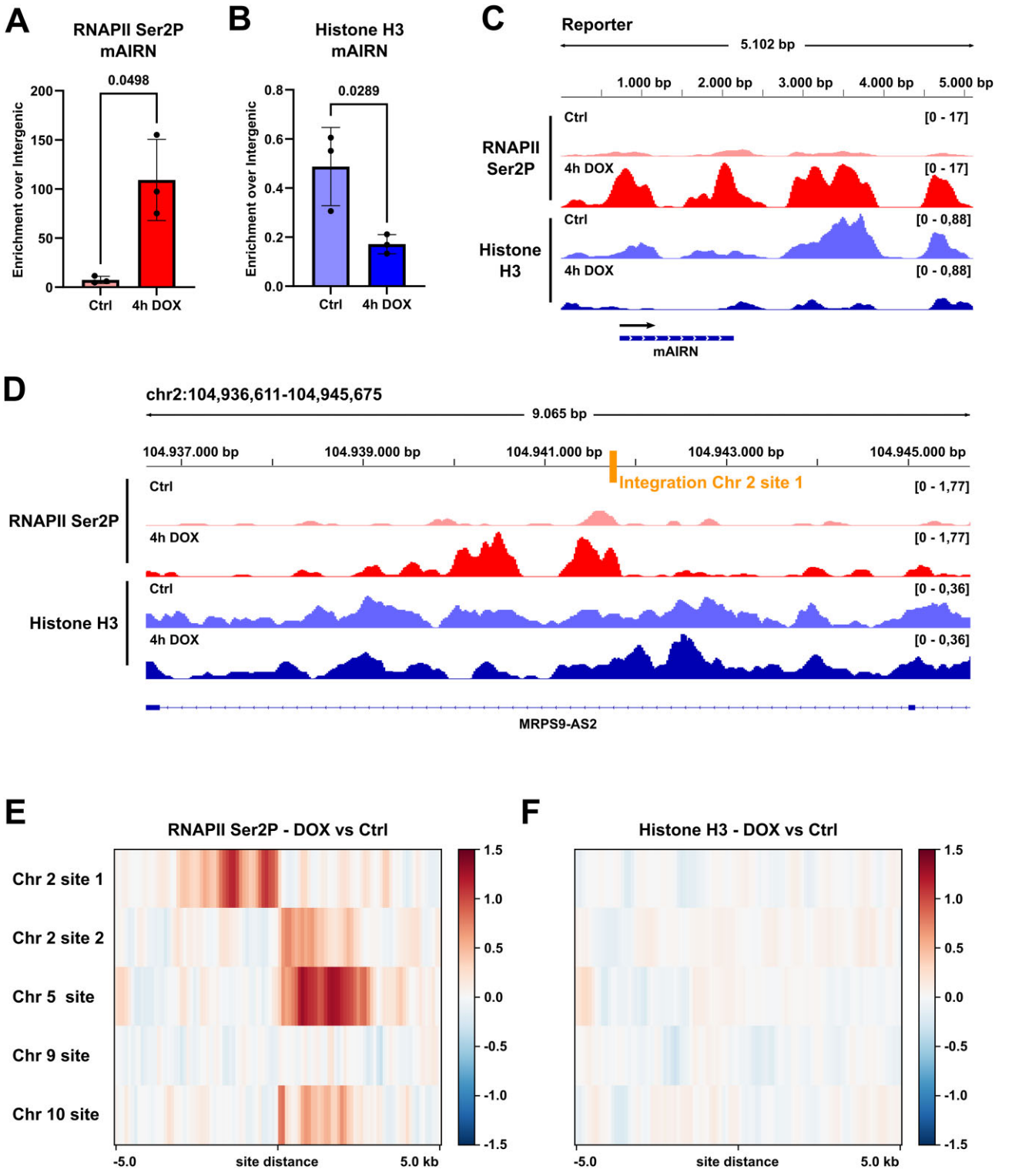


Figure 4. TRC induction disrupts local chromatin structure on individual integration sites. (**A**) ChIP-qPCR analysis showing RNAPII Ser2P levels at the reporter site (mAIRN#1 primers) in synchronized S-phase cells 4 h after release from double thymidine block. Cells were treated with 0 or 1 μg/mL DOX for 4 h ($n = 3$). Error bars indicate SD. Welch's t-test (**B**) ChIP-qPCR analysis showing histone H3 levels at the reporter site (mAIRN#1 primers) in the same conditions as (A). (**C**) Representative genome browser snapshot showing RNAPII Ser2P and histone H3 occupancy across the entire reporter construct in DOX treated or untreated conditions in synchronized S-phase cells. Sequencing libraries were derived from samples shown in (A) and (B). (**D**) Representative genome browser snapshot showing RNAPII Ser2P and histone H3 occupancy at the MRPS9-AS2 locus which contains the Chr 2 site 1 integration site (exact position highlighted with a bar) in DOX treated or untreated conditions in synchronized S-phase cells. Sequencing libraries were derived from samples shown in (A) and (B). (**E**) Heatmap showing log2 fold change of RNAPII Ser2P signal upon DOX treatment over control in a ± 5 kb region around the integration site locations, 100 bp bin size. (**F**) Heatmap showing log2 fold change of histone H3 signal upon DOX treatment over control in a ± 5 kb region around the integration site locations, 100 bp bin size.

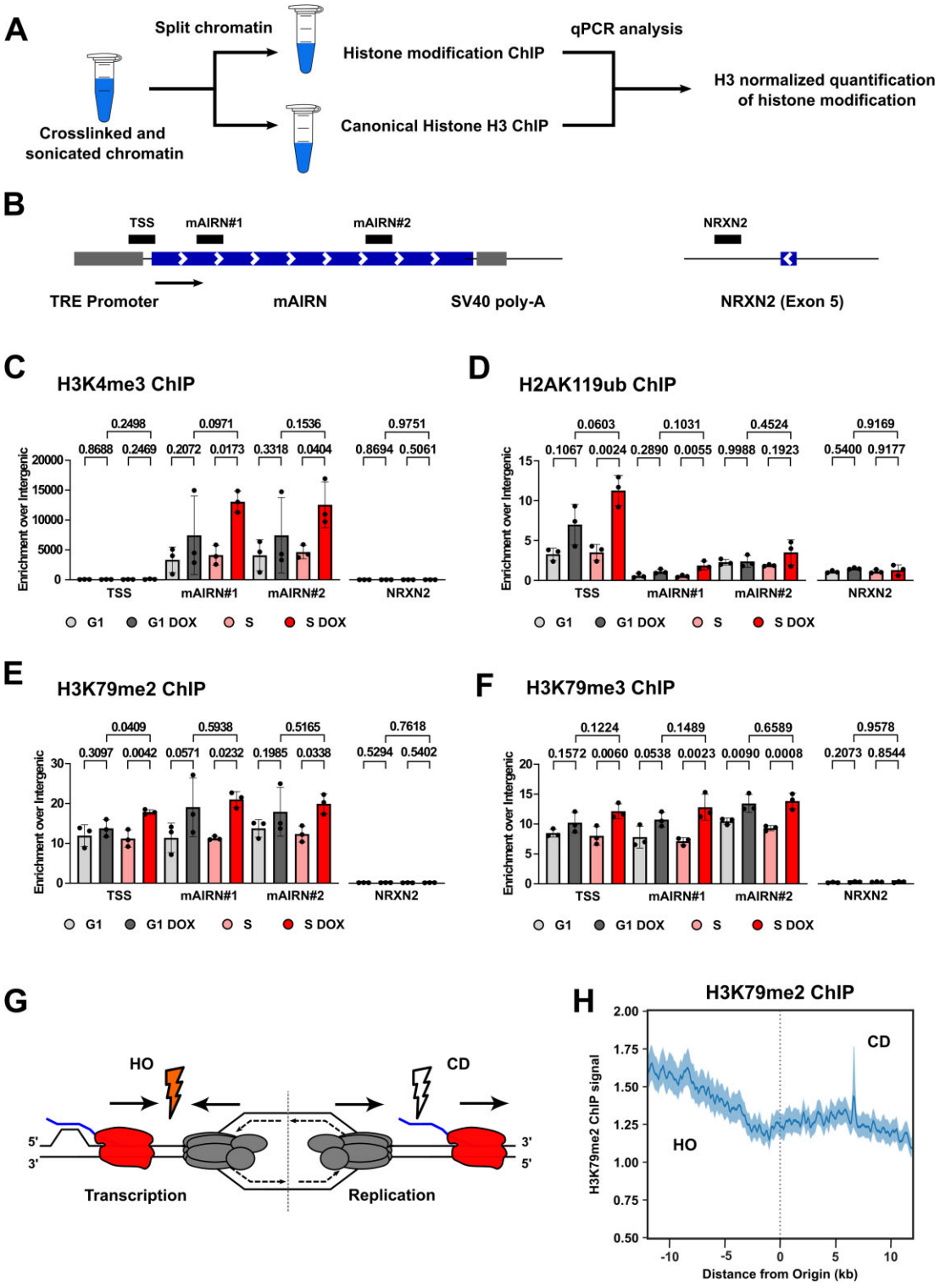


Figure 5. H3K79 methylation is a TRC-enriched chromatin modification at the R-loop reporter and genome-wide. (**A**) Schematic representation of the ChIP workflow used for screening TRC-dependent histone modifications including canonical histone H3 normalization. (**B**) Schematic representation of the mAIRN region of the integrated reporter construct and control region near NRXN2 exon 5. Locations of the tested primer pairs (TSS, mAIRN#1, mAIRN#2, NRXN2) in subsequent ChIP experiments are highlighted with bars above. (**C**) H3 normalized H3K4me3 ChIP in G1 or 4 h released S-phase cells treated with 0 or 1 μg/mL DOX. H3K4me3 levels were tested at the reporter sequence with primers at TSS, mAIRN#1, mAIRN#2, or the NRXN2 control site ($n = 3$). Error bars indicate SD. Ordinary one-way ANOVA with Tukey's multiple comparison test. (**D**) H3 normalized H2AK119ub ChIP in the same conditions as (C). (**E**) H3 normalized H3K79me2 ChIP in the same conditions as (C). (**F**) H3 normalized H3K79me3 ChIP in the same conditions as (C). (**G**) Cartoon for the selection of genomic regions biased toward HO versus CD collisions by the identification of intragenic origins of replication within actively transcribed genes [7]. (**H**) Analysis of H3K79me2 ChIP signal from HeLa cells at intragenic origins within actively transcribed genes. The analysis windows around the regions are 24 kb in size and excluded from the analysis if positioned within 5 kb from promoters and terminators. H3K79me2 signal accumulates at HO side of origins in gene bodies compared to the CD side. Error bands represent a 95% confidence interval as determined by a bootstrap of the mean.

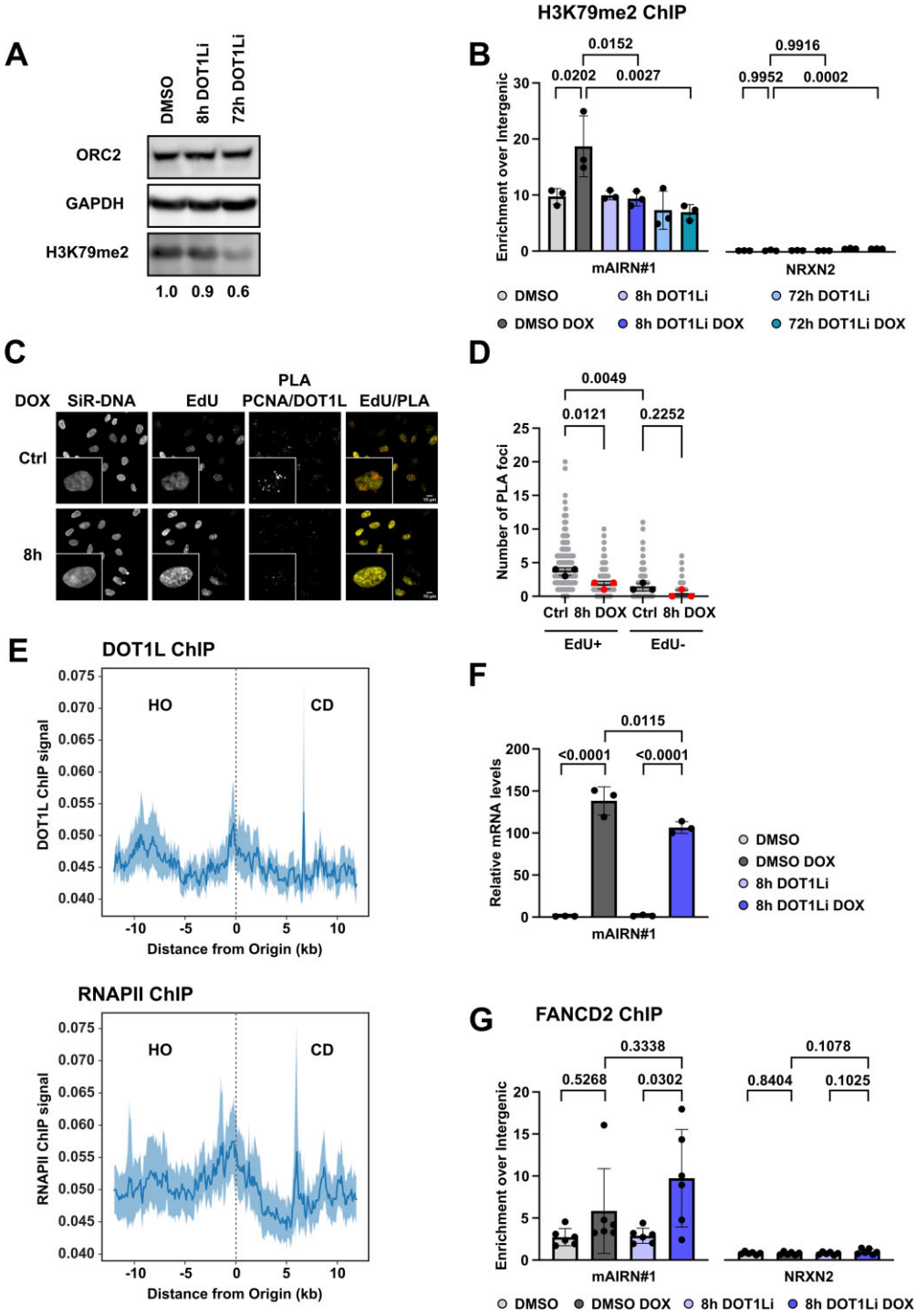


Figure 6. Evaluation of the role of H3K79 methylation at TRC sites. (**A**) Representative Western blot of H3K79me2 upon 5 μM DOT1L inhibition (EPZ-5676) or DMSO control treatment for 8 or 72 h. GAPDH and ORC2 loading controls. Quantifications of H3K79me2 signal relative to the DMSO condition shown below. (**B**) H3 normalized H3K79me2 ChIP in cells treated with 5 μM DOT1L inhibition (EPZ-5676) for 8 and 72 h as well as 0 or 1 μg/mL DOX. H3K79me2 levels were tested at the mAIRN#1 or the NRXN2 control site ($n = 3$). Error bars indicate SD. Ordinary one-way ANOVA with Tukey's multiple comparison test. (**C**) Representative images of PLA assay with DOT1L and PCNA antibodies. EdU Click-it staining was performed to label S-phase cells. Cells were treated with 0 or 1 μg/mL DOX for TRC induction. Scale bar 10 μm. (**D**) Quantification of (C) in EdU positive and negative cells ($n = 3$). Error bars indicate SD. Ordinary one-way ANOVA with Tukey's multiple comparison test. (**E**) Analysis of DOT1L and RNAPII ChIP signal from MOLM13 cells at intragenic origins within actively transcribed genes previously defined in HeLa cells. The analysis windows around the regions are 24 kb in size and excluded from the analysis if positioned within 5 kb from promoters and terminators. DOT1L and RNAPII signal overlaps and accumulates at HO side of origins in gene bodies compared to the CD side. Error bands represent a 95% confidence interval as determined by a bootstrap of the mean. (**F**) RT-qPCR analysis of mAIRN RNA expression using primer pair mAIRN#1 in cells treated with 5 μM DOT1L inhibition (EPZ-5676) or DMSO control treatment. Additionally, 0 or 1 μg/mL DOX were added for 4 h ($n = 3$). Error bars indicate mean values with SDs. Ordinary one-way ANOVA with Tukey's multiple comparison test. (**G**) FANCD2 ChIP using primer pair mAIRN#1 or NRXN2 in cells treated with 5 μM DOT1L inhibition (EPZ-5676) or DMSO control treatment. Additionally, 0 or 1 μg/mL DOX were added for 24 h ($n = 3$). Error bars indicate SDs. Ordinary one-way ANOVA with Tukey's multiple comparison test.

the TSS and 5′ end of the gene (Fig. 5D, TSS and mAIRN#1), whereas less or no differential behavior between conditions was observed at the mAIRN 3′ end and NRXN2 gene (Fig. 5D, mAIRN#2 and NRXN2). Finally, we also monitored the dynamic changes of H3K79 di- and trimethylation (H3K79me2, me3) upon TRC induction at the mAIRN reporter. H3K79 methylation was not only suggested as an activating mark enriched at both promoters and gene bodies of actively transcribed genes [76,77] but also important for effective DNA damage repair [78,79]. Strikingly, our induced TRC reporter gene accumulated both H3K79me2 (Fig. 5E) and H3K79me3 (Fig. 5F) across all tested mAIRN primer locations particularly in S-phase released cells, whereas no change was observed at the NRXN2 control locus. To further corroborate these results, we gathered ENCODE ChIP-seq data of multiple histone modifications and overlapped their enrichment with genomic regions biased toward HO versus CD collisions as previously shown (Fig. 5G and [7]). In contrast to H3K4me3, H3K27me3, and H3K36me3 which showed a symmetric distribution (Supplementary Fig. S6C–E), H3K79me2 and H3K79me3 showed a significant enrichment at the HO region compared to CD (Fig. 5H, Supplementary Fig S6F), suggesting that this modification could specifically mark HO collision sites. Consistently, H3K79me2 and H3K79me3 ChIP-Seq data showed a notable overlap with R-loop positive regions as determined by qDRIP-Seq (Supplementary Fig. S6G and H). Using the same approach, we also tested the distribution of H2AK119ub over these TRC-prone regions but could not observe a meaningful enrichment of this histone mark over these TRC-prone regions (Supplementary Fig. S6I). This may be attributed to the fact that no publicly available H2AK119ub ChIP-Seq dataset was available in HeLa cells, but only in MCF7 breast cancer cells, thus making this analysis across different cell lines less robust. Nevertheless, these findings suggest that our TRC reporter cell line can successfully recapitulate the dynamic changes of previously identified TRC-related chromatin modifications such as H3K4me3 and H2AK119ub and supports the notion that H3K79 methylation is also a modification relevant to TRC and R-loop biology.

Evaluation of the role of H3K79 methylation at TRC sites

The only known histone methyltransferase catalyzing H3K79 methylation is the disruptor of telomeric silencing-1-like (DOT1L) [77, 80]. To further investigate the function of DOT1L-mediated H3K79 methylation deposition at TRC sites, we aimed to inhibit DOT1L activity in our TRC reporter cell line. For this, we took advantage of the small molecule inhibitor EPZ-5676 (Pinometostat, DOT1Li), which was previously shown to specifically inhibit DOT1L methyltransferase activity [81]. Western blot analysis of global H3K79me3 levels showed that 1 μM or 5 μM DOT1Li treatment resulted in a global loss of H3K79me3 with prolonged (72 h) treatment, but not after acute (8 h) inhibition treatment compared to GAPDH or ORC2 as a nuclear loading control (Fig. 6A, Supplementary Fig. S7A). This result suggests that the drug is active and inhibits new H3K79 methylation deposition, but it takes 2–3 cell doublings to dilute the pre-existing mark from total chromatin, consistent with previous studies [78, 81]. Next, we assessed by ChIP-qPCR whether H3K79me2 deposition was also specifically impaired at the mAIRN gene in our TRC reporter cell line. As expected, DMSO-treated cells showed a clear DOX-dependent increase in H3K79me2 levels at mAIRN (Fig. 6B), but not at unrelated NRXN2 (Fig. 6B) or ACTB control genes (Supplementary Fig. S6B). Strikingly, both acute (8 h) and prolonged (72 h) DOT1Li treatment completely abolished the DOX-dependent increase of H3K79me2 (Fig. 6B), suggesting that DOT1L is specifically activated to methylate H3K79 at our mAIRN integration sites upon DOX induction. To test the possibility whether DOT1L mediated H3K79 methylation could affect nucleosome stability and thus promote the observed nucleosome loss, we have tested the effect of DOT1L inhibition on H3 levels at the mAIRN reporter sequence. DOT1L inhibition for 8h as well as 72 h did not prevent or reduce the nucleosome loss phenotype at the mAIRN sites upon DOX induction, suggesting that DOT1L mediated H3K79 methylation is not required for nucleosome eviction (Supplementary Fig. S7C).

To obtain more evidence that DOT1L activity is connected to TRCs, we tested whether DOT1Li can exacerbate PLA foci number between RNAPIIpS2 and PCNA. Consistent with our previous data (Fig. 2E), challenging cells with DOX induction for 4 h had little impact on TRC levels in the presence of DOT1L activity. Interestingly, DOT1Li for 72 h reduced the overall TRC burden in the basal state without TRC induction, which is likely the result of global transcription impairment and resulting reduced levels of RNAPIIpS2 on chromatin as previously reported [82]. Strikingly, upon 4 h of DOX induction, addition of DOT1Li for 72 h led to an increase in TRC PLA foci, suggesting that DOT1L inhibition impairs the cellular T-R coordination and the ability to overcome conflicts (Supplementary Fig. S7D). To obtain further functional insights, we considered the possibility that DOT1L may directly associate with active replication forks. Indeed, a significant number of PCNA-DOT1L PLA foci could be detected in EdU-positive S-phase cells. To confirm specificity of this PLA combination, we used siRNA-mediated knockdown of DOT1L (Supplementary Fig. S7E) and observed significant reduction in PCNA-DOT1L-PLA foci specifically in S-phase cells compared to siControl cells (Supplementary Fig. S7F). Interestingly, DOX-dependent TRC induction resulted in a reduction of DOT1L-PCNA PLA foci in S-phase cells (Fig. 6C and D), suggesting that DOT1L can dissociate from ongoing replication forks, thereby potentially acting on post-replicative chromatin. Consistently, two recent studies showed that RNAPII complexes can quickly re-associate with active genes on post-replicative chromatin and rapidly resume transcription [28,29]. To test the possibility that DOT1L-mediated H3K79 methylation activity is required for this transcriptional restart at TRC sites, we gathered available RNAPII, H3K79me2 and DOT1L ChIP-Seq data from a leukemia cell line and overlapped their enrichment with HO and CD TRC sites [7] as shown in Fig. 5H. H3K79me2 also specifically marked HO collision sites in this cell line as observed before in HeLa cells (Supplementary Fig. S7G). Strikingly, RNAPII and DOT1L also showed a significant enrichment and overlap at HO collision sites (Fig. 6E), implicating DOT1L activity in proficient RNAPII elongation at TRC sites. Consistently, 8 h DOT1Li treatment also significantly reduced transcriptional output from the mAIRN gene as observed by RT-qPCR (Fig. 6F), but not at the unrelated PUM1 and ALAS1 housekeeping genes (Supplementary Fig. S7H). To confirm specificity of these data to DOT1L activity, we repeated the experiment with a different, structurally distinct DOT1L inhibitor (EPZ004777) and found the same small but significant decrease in mAIRN tran-

scriptional output as with EPZ-5676 after short-term (8 h) DOT1L inhibition (Supplementary Fig. S7I and J). In addition, we tested the effect of RNAi-mediated depletion of DOT1L on mAIRN transcription levels. Although, we only achieved a ∼50% reduction of DOT1L and H3K79me2 levels after 72 h knockdown (Supplementary Fig. S7E), we observed consistent with the chemical inhibition experiments a significant decrease in mAIRN transcriptional output upon DOX treatment (Supplementary Fig. S7K), but not at the unrelated PUM1 and ALAS1 housekeeping genes (Supplementary Fig. S7L). Together, this provides further evidence that the full transcriptional potential of mAIRN after TRC induction is dependent on DOT1L enzyme activity.

Finally, we asked whether the presence of H3K79 methylation has functional relevance and mitigates potential DNA damage at our TRC sites. To this end, we performed FANCD2 ChIP after TRC induction with and without 8 h DOT1Li treatment and found that FANCD2 recruitment is further increased upon DOT1L inhibition at the activated mAIRN reporter but not the NRXN2 control locus (Fig. 6G), suggesting that the DNA damage inflicted at the TRC site is exacerbated without active H3K79 methylation. Together, these data support a model (Fig. 7) where DOT1L activity is required for TRC resolution by restoring transcriptional processivity of the RNAPII complex, thereby preventing TRCs and elevated DNA damage.

Discussion

We have established a cell-based system to reliably induce TRCs at defined chromosomal sites and used it to characterize the impact of TRCs on the underlying local chromatin landscape. Unexpectedly, we observed that TRCs lead to a loss of nucleosome occupancy at the reporter sites that is strongly correlated with local R-loop formation and impaired DNA replication fork progression. This results in locus-specific DNA damage as indicated by elevated FANCD2 occupancy at the reporter, but also leads to a global DNA damage signaling response as indicated by elevated FANCD2 foci formation and γH2AX signal across the nucleus (Fig. 2). Furthermore, screening of multiple TRC-associated histone modifications confirmed the enrichment of previously identified TRC-related chromatin modifications and uncovered H3K79 methylation as an important chromatin mark enriched at TRC sites. Preventing deposition of new H3K79me2/3 at TRC sites by inhibiting its methyltransferase DOT1L exacerbates the DNA damage response and prevents productive transcription from the conflict site (Fig. 7). Overall, our results provide mechanistic insight into how TRCs and associated R-loops threaten chromatin structure by disrupting nucleosome organization and DNA replication and define H3K79 methylation as an important histone modification at chromosomal TRC sites.

Stable R-loop formation at TRC sites is incompatible with nucleosome incorporation

A consistent result from our multiple approaches and TRC model systems is that stable R-loop formation at the mAIRN sequence is incompatible with nucleosome assembly on the three-stranded nucleic acid structure. First, *in vitro* reconstitution of mAIRN RNA:DNA hybrids demonstrates that RNA:DNA hybrids are resistant to nucleosome formation in competition with dsDNA. Second, HO-TRC, but not the corresponding CD-TRC plasmids, display a substantial reduction of nucleosomes, as measured in independent MNase and ChIP-qPCR assays (Fig. 1D and E), consistent with our previous finding that only HO-TRCs form stable R-loops on the plasmid reporter system [7]. Third, the chromosomally integrated mAIRN sequence is also prone to loss of nucleosomes upon transcription activation in both G1 and S-phase cells (Fig. 4B, Supplementary Fig. S6A), indicating that R-loop formation alone is sufficient to displace nucleosomes. Importantly, S-phase-released cells show a more pronounced nucleosome loss that correlates with a higher R-loop burden (Fig. 3B), suggesting that TRC-stabilized R-loops have an enhanced ability to disrupt nucleosome organization. It is also important to note that nucleosome eviction could only be observed at the R-loop mediated TRC reporter sites but not at neighboring regions to the integration site that show clearly upregulated transcription (Fig. 4E and F), making a strong argument that the disruption of nucleosomes is not a simple consequence of transcription activation. In agreement, we could not observe nucleosome eviction in a control cell line with integrated non-R-loop forming ECFP sequences (Supplementary Fig. S4F and G), providing independent evidence that the observed phenotype is specific to the R-loop forming mAIRN TRC reporter loci. Together, we conclude that R-loop formation is a potent driver of nucleosome disruption, thus leaving the underlying genomic region in an open, accessible state. This is fully consistent with the early observation that the presence of RNA in a double-helical DNA is inhibiting its interaction with histone proteins [83]. Based on a subsequent structural characterization of R-loops, RNA:DNA hybrids adopt an intermediate form between A-DNA and B-DNA [64]. This unusual secondary structure shows less flexibility and is thus unlikely to wrap efficiently around the histone octamers, providing a plausible explanation for the observed nucleosome loss phenotypes. Similar to exposure of ssDNA in the non-template strand [84], nucleosome depletion at R-loops could offer another layer of explanation for R-loop and TRC-driven genomic instability since nucleosome-free genomic regions are more prone to DNA damage [85,86].

In summary, our work establishes an intertwined relationship between R-loop formation, TRCs, and chromatin disruption. It also underscores the importance of proper R-loop resolution and removal to maintain genome and epigenome integrity. Future investigations using global R-loop stabilization approaches, such as RNase H1/2 inactivation, will provide crucial insights into the genome-wide relationship between R-loops, TRC occurrence and nucleosome organization.

Genomic integration of an inducible mAIRN R-loop sequence allows locus-specific TRC analysis

Previous work in numerous prokaryotic and eukaryotic model systems has defined TRCs as a threat to genome integrity and a cause of mutations [7, 9, 11, 87, 88]. Crucially, their stochastic occurrence and dependence on two highly complex nuclear processes make TRCs difficult to study. Hence, researchers have frequently employed plasmid-based reporter systems in bacteria, yeast, and human cells to study TRCs under controlled conditions [7, 9, 11]. Such reporter systems are complemented by correlative studies which rely on drug treatments or oncogene activation to induce TRCs in a genome-wide manner and then correlate co-occurring DNA damage responses to TRCs [71, 89, 90]. Although useful, plasmid-based reporter systems likely do not capture the full complex-

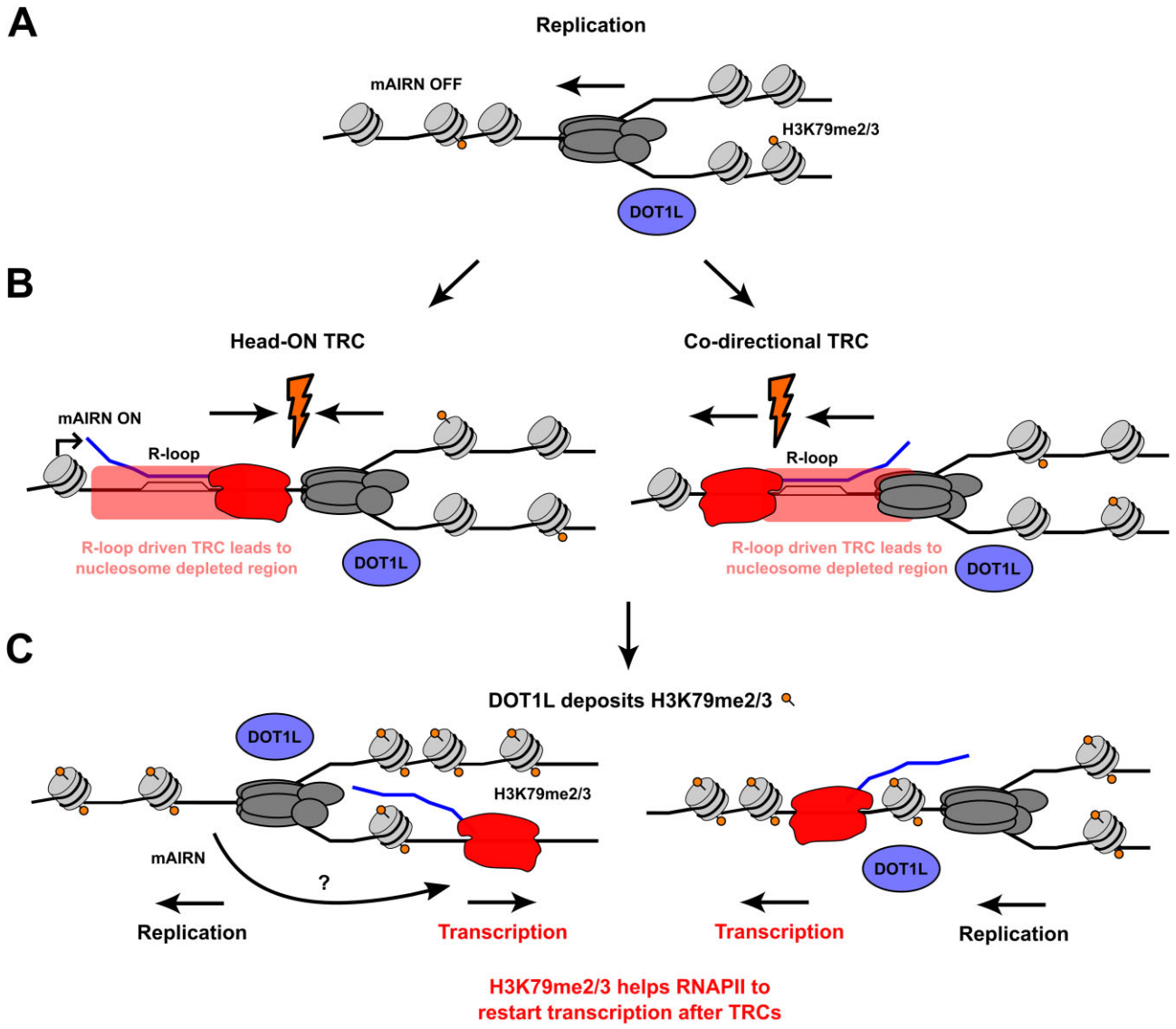


Figure 7. Model for the functional role of H3K79 methylation at TRC sites to allow effective transcription recovery. (**A**) Cells not transcribing the mAIRN reporter (OFF) can replicate normally without disruption of chromatin organization. (**B**) Upon transcriptional activation of the mAIRN reporter (ON), the mAIRN reporter forms R-loops and interferes with replication fork progression. The resulting HO or CD TRC with an associated R-loop causes a local reduction of nucleosome occupancy. Simultaneously, DOT1L dissociates from the replication machinery and deposits H3K79me2/3. (**C**) H3K79me2/3 helps to efficiently recover transcription at the mAIRN reporter site.

ity and dynamics of TRC induction and resolution on endogenous chromatin, whereas global TRC induction approaches are prone to systemic cellular responses and thus secondary side-effects not related to individual TRC events. Altogether, this prevents true mechanistic insights how TRCs impact chromatin structure. In the present study, we constructed a TRC reporter system based on chromosomal integration of the R-loop forming mAIRN sequence previously used in the human plasmid reporters [7]. We chose not to integrate the unidirectional oriP/EBNA1 replication origin due to late and inefficient replication activity of this origin in human cells. Consequently, endogenous origins would likely replicate the integrated sequence before oriP could become active [66, 67] and remove the orientational advantage from the plasmid system. R-loop formation has been shown to impair RNAPII progression [91]. We speculated that R-loop-driven RNAPII slowdown at mAIRN sequences would create obstacles for incoming replication forks, thus inducing TRCs. Integration of the reporter construct with Sleeping Beauty transposase at multiple random locations (Fig. 2B) also has the advantage of detecting the averaged signals from different genomic environments and chromatin contexts, reducing potential locus-specific biases. Thus, our chromosomal reporter provides a controlled environment for TRC induction while simultaneously residing in the endogenous chromatin and will thus serve as a useful tool for future mechanistic studies of TRCs and their relationships with R-loop biology, TRC-driven mutation burden, and (epi)genome instability.

H3K79 methylation is deposited at TRC sites and provides a chromatin environment for effective transcription recovery

Various histone modifications such as H3K4me3 and H2AK119ub have been recently connected with genomic TRC regions [32, 33]. As we could also observe an enrichment of

these histone PTMs in our TRC reporter cell line (Fig. 5), this suggests that our approach can successfully recapitulate the dynamic chromatin changes of endogenous TRC sites. Interestingly, we also observed specific enrichment of H3K79me2 and H3K79me3 at our TRC reporter site as well as other genomic sites prone to R-loop forming HO collisions (Fig. 5H). H3K79me2/3 is deposited by DOT1L, the sole enzyme responsible for this modification in mammalian cells. While known to be associated with transcription elongation and found in various transcription elongation complexes [77], the exact function of H3K79me2/3 in transcription regulation remains unclear. Additionally, H3K79me2/3 has been linked to DNA damage response, as the 53BP1 Tudor domain binds H3K79me3 [92], which is crucial for responding to double-strand DNA breaks. Furthermore, H3K79me2/3 interacts with PCNA [79], suggesting a role in replication fork elongation. Considering these links to both transcription and replication processes, a connection with TRCs appears likely but has not been directly investigated. Indeed, our study shows that H3K79me2 is actively deposited at the induced TRC reporter (Fig. 6B), suggesting a dynamic rather than a stable modification in this context. This response might differ from the global role of H3K79me2/3 in aiding RNAPII processivity at active genes. Notably, inhibiting H3K79me2/3 deposition at the TRC reporter leads to reduced mAIRN RNA output. Furthermore, high levels of H3K79me2/3 are found in genes prone to forming R-loops (Supplementary Fig. S6F and G), suggesting a potential role of this modification in maintaining or restarting transcription at genes prone to R-loop dependent replication fork stalling. Recently, it was shown that RNAPII transcription recovery at TRC sites is mediated by the transcription elongation factor ELL [16], and it will be very interesting in future studies to dissect whether or not DOT1L and ELL collaborate in the same pathway. In conclusion, our study sheds light on the role of H3K79me2/3 in TRCs and likely acts as a platform to facilitate TRC resolution by aiding the restart of transcription after collisions between replication and transcription machineries.

TRCs impair DNA replication fork progression and sensitize the cells to ATR kinase inhibition

TRCs and R-loops have been shown to be obstacles to DNA replication fork progression in different model systems [21, 38]. The chromosomal reporter system enabled us to demonstrate local TRC-driven DNA replication progression impairments in endogenous mammalian chromatin. Crucially, impaired DNA replication at the induced TRCs occurred at the expected replication timing of the superordinate replication domain (Fig. 3D–F, Supplementary Fig. S3F), showing that TRCs impair replication locally but cause no global DNA replication timing disruption (Supplementary Fig. S3E). As the majority of reporter sites in the Clone#12 cell line reside within early replicating regions of the genome (Fig. 3D), the observed replication delay at the TRC reporter genes was strongest at the earliest 2 h S-phase time point (Fig. 3E and F, Supplementary Fig. S3F), but eventually replicated after 4 and 6 h S-phase progression (Supplementary Fig. S3G). This paucity of genome replication in a later stage of the cell cycle is reminiscent of the recently identified G2/M DNA synthesis (G-MiDS) sites that are highly transcribed TSS during replication and require RNAPII removal in G2/M phase to complete DNA synthesis [42]. However, G-MiDS sites are much smaller up to a few hundred basepairs and not associated with increased DNA damage, which is in contrast to our reporter sites that are much larger (∼ 5 kb), resolved earlier during S-phase progression, and cause local and global DNA damage phenotypes (Fig. 2I and J). This suggests that the TRC resolution at the mAIRN reporter during S-phase is fundamentally different from G-MiDS and follows a pathway detected by the DNA damage checkpoint. Consistently, we show that persistent DOX induction over several cell divisions causes a mild decrease in cell proliferation (Fig. 2K and L, Supplementary Fig. S2P and Q). Crucially, this proliferation defect is strongly exacerbated upon treatment with an ATR kinase inhibitor, emphasizing the importance of ATR signaling in the cellular TRC response. Our previous work on the plasmid system has shown that HO conflicts induce ATR activation, while CD conflicts induce ATM [7]. Curiously, ATM kinase inhibition did not show strong synergistic effects with TRC induction (Fig. 2L, Supplementary Fig. S2Q). Thus, it is tempting to speculate that unresolved HO conflicts are the main drivers of ATR inhibition-driven sensitivity in our TRC reporter cells. ATR inhibitors have shown promising results as anti-cancer drugs in clinical trials [93]. Thus, stratifying patient samples with high or low TRC burden could thereby sensitize and further improve the efficiency of ATR inhibitor application in cancer treatment.

mAIRN R-loop formation as a model system for TRC and replication stress induction in endogenous chromatin

Although cells inducing the non R-loop forming ECFP control gene did not display reduced nucleosome occupancy or fork stalling phenotypes (Supplementary Figs. S2K, S4F and G), it is important to note that we were unable to rescue the observed H3 loss and accumulation of FANCD2 at induced mAIRN loci by overexpression of RNase H1 (Supplementary Fig. S4D and E), making the direct R-loop dependency of these effects in our cell line a matter of debate. In our experimental setup, we maintained the cells for 24 h in the presence of DOX allowing for continuous transcription/R-loop formation and RNase H1 overexpression, which may result in a higher turnover between R-loop formation and degradation but not suffice to substantially reduce their levels in the cell population. In addition, the mAIRN sequence has a very high GC-skew, a feature that is highly correlated with R-loop formation but also a feature of previously reported RNaseH-resistant R-loops [60] that are highly confined to regions with high GC-skew. Finally, RNase H1 overexpression also resulted in cell toxicity and cell death, complicating the interpretation of our experiments as we cannot rule out whether the surviving cells are the ones not expressing enough RNase H1 or whether they have been rewired to cope with an excess of RNase H1 protein.

Besides these more technical explanations why our RNase H1 overexpression construct was unable to efficiently degrade and/or remove the particularly strong and stable mAIRN R-loops, an alternative hypothesis is that the replication stress phenotype is caused by distinct non-R-loop DNA secondary structures that have been previously described to form at the mAIRN non-template strand *in vitro* in an RNase H-resistant manner [94]. To distinguish between these two possibilities, an interesting direction for the future will be to introduce specific mutations in our integrated mAIRN copies that

were previously shown to prevent the formation of these additional mAIRN secondary structures but not interfere with RNA:DNA hybrid formation.

Finally, it will be important to validate our findings at the integrated mAIRN reporter sites at endogenous TRCs occurring at R-loop prone loci. To this end, a recent study has explored the chromatin landscape of putative genomic TRC regions in K562 cells by combining DRIPc-seq as marker of R-loop prone sites with OK-seq to define replication fork directionality and FANCD2 ChIP-seq to identify sites of fork stalling [31]. It is worth mentioning that overlapping these regions with ENCODE ChIP-seq data showed that H3K79me2 represents one of the most enriched chromatin marks at such putative genomic TRC sites [31], thereby independently validating one of the key results from our TRC reporter cell line. Nevertheless, binarizing the genome in TRC and non-TRC regions might be overly simplistic since genes are replicated at most 80:20 in one direction versus the other and transcription initiation occurs in bursts that fluctuate between an ON and OFF state in a stochastic manner. Therefore, it is impossible from such bulk and correlative analyses of TRC-associated markers to unequivocally conclude that a TRC is indeed occurring in an individual genomic locus of an individual cell. To address this, novel methodology will need to be developed by the community to directly and experimentally map TRCs. We consider the development of the TRC reporter cell line described in this study here as a first step toward this goal.

Acknowledgements

We thank all past and current Hamperl lab members, Inti de la Rosa Velazquez from the Sequencing Facility at Helmholtz Munich and Andreas Ettinger from the IES Imaging Core Facility, Mrinmoy Pal for help and support with high-throughput sequencing data management, Joshua Saldivar, and Maria-Elena Torres-Padilla for helpful suggestions, feedback and critical reading of the manuscript. We thank Kristina Schreiber for the *in vitro* nucleosome reconstitution experiments. We thank Elisabeth Kruse for expert technical support and Michael Kirsch for providing us with the U-2 OS Tet-ON parental cell line. We thank Julian Stingele and lab members for the opportunity to use their Incucyte S3 Live-Cell Analysis System.

Author Contributions: Conceptualization: S.H.; Data Curation: M.W., M.T., T.S., and E.M.G.; Formal Analysis: M.W., M.T., T.S., and E.M.G.; Funding acquisition: S.H.; Investigation: Everyone; Supervision: S.H.; Validation: Everyone; Writing – original draft: M.W. and S.H.; Writing – review and editing: Everyone.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

Work in the Hamperl laboratory is funded by the Helmholtz Association, the German Research Foundation (DFG) Project-ID 213249687 (SFB 1064), and the European Research Council (ERC starting grant 852798). Funding to pay the Open Access publication charges for this article was provided by H2020 European Research Council.

Data availability

BrdU-Seq, ChIP-Seq, and Whole Genome Sequencing data are available on GEO with the accession number GSE267494.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

References

1. Zeman MK, Cimprich KA. Causes and consequences of replication stress. *Nat Cell Biol* 2014;**16**:2–9. https://doi.org/10.1038/ncb2897
2. López-Otín C, Blasco MA, Partridge L *et al.* The hallmarks of aging. *Cell* 2013;**153**:1194–217. https://doi.org/10.1016/j.cell.2013.05.039
3. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;**100**:57–70.
4. García-Muse T, Aguilera A. Transcription–replication conflicts: how they occur and how they are resolved. *Nat Rev Mol Cell Biol* 2016;**17**:553–63.
5. Lalonde M, Trauner M, Werner M *et al.* Consequences and resolution of transcription-replication conflicts. *Life (Basel, Switzerland)* 2021;**11**:637.
6. Goehring L, Huang TT, Smith DJ. Transcription–replication conflicts as a source of genome instability. *Annu Rev Genet* 2023;**57**:157–79.
7. Hamperl S, Bocek MJ, Saldivar JC *et al.* Transcription-replication conflict orientation modulates R-loop levels and activates distinct DNA damage responses. *Cell* 2017;**170**:774–86. https://doi.org/10.1016/j.cell.2017.07.043
8. Merrikh H. Spatial and temporal control of evolution through replication-transcription conflicts. *Trends Microbiol* 2017;**25**:515–21. https://doi.org/10.1016/j.tim.2017.01.008
9. Lang KS, Hall AN, Merrikh CN *et al.* Replication-transcription conflicts generate r-loops that orchestrate bacterial stress survival and pathogenesis. *Cell* 2017;**170**:787–99.
10. Kim N, Abdulovic AL, Gealy R *et al.* Transcription-associated mutagenesis in yeast is directly proportional to the level of gene expression and influenced by the direction of DNA replication. *DNA Repair (Amst)* 2007;**6**:1285–96. https://doi.org/10.1016/j.dnarep.2007.02.023
11. Prado F, Aguilera A. Impairment of replication fork progression mediates RNA polII transcription-associated recombination. *EMBO J* 2005;**24**:1267–76. https://doi.org/10.1038/sj.emboj.7600602
12. Petermann E, Lan L, Zou L. Sources, resolution and physiological relevance of R-loops and RNA–DNA hybrids. *Nat Rev Mol Cell Biol* 2022;**23**:521–40. https://doi.org/10.1038/s41580-022-00474-x
13. Brickner JR, Garzon JL, Cimprich KA. Walking a tightrope: The complex balancing act of R-loops in genome stability. *Mol Cell* 2022;**82**:2267–97. https://doi.org/10.1016/j.molcel.2022.04.014
14. Kumar C, Remus D. Looping out of control: R-loops in transcription-replication conflict. *Chromosom* 2023;**133**:37–56.
15. Kumar C, Batra S, Griffith JD *et al.* The interplay of RNA:DNA hybrid structure and G-quadruplexes determines the outcome of R-loop-replisome collisions. *eLife* 2021;**10**:e72286. https://doi.org/10.7554/eLife.72286
16. Chappidi N, Nascakova Z, Boleslavska B *et al.* Fork cleavage-religation cycle and active transcription mediate replication restart after fork stalling at co-transcriptional R-loops. *Mol Cell* 2020;**77**:528–41. https://doi.org/10.1016/j.molcel.2019.10.026
17. Zardoni L, Nardini E, Brambati A *et al.* Elongating RNA polymerase II and RNA:DNA hybrids hinder fork progression and

gene expression at sites of head-on replication-transcription collisions. *Nucleic Acids Res* 2021;**49**:12769–84. https://doi.org/10.1093/nar/gkab1146
18. Aiello U, Challal D, Wentzinger G *et al.* Sen1 is a key regulator of transcription-driven conflicts. *Mol Cell* 2022;**82**:2952–66. https://doi.org/10.1016/j.molcel.2022.06.021
19. Barroso S, Herrera-Moyano E, Muñoz S *et al.* The DNA damage response acts as a safeguard against harmful DNA-RNA hybrids of different origins. *EMBO Rep* 2019;**20**:e47250.https://doi.org/10.15252/embr.201847250
20. Šviković S, Crisp A, Tan-Wong SM *et al.* R-loop formation during S phase is restricted by PrimPol-mediated repriming. *EMBO J* 2019;**38**:e99793.https://doi.org/10.15252/embj.201899793
21. Stoy H, Zwicky K, Kuster D *et al.* Direct visualization of transcription-replication conflicts reveals post-replicative DNA:RNA hybrids. *Nat Struct Mol Biol* 2023;**30**:348–59. https://doi.org/10.1038/s41594-023-00928-6
22. Heuz J, Kemiha S, Barthe A *et al.* RNase H2 degrades toxic RNA:DNA hybrids behind stalled forks to promote replication restart. *EMBO J* 2023;**42**:e113104. https://doi.org/10.15252/EMBJ.2022113104
23. Gómez-González B, Aguilera A. Transcription-mediated replication hindrance: a major driver of genome instability. *Genes Dev* 2019;**33**:1008–26.
24. Brüning JG, Marians KJ. Replisome bypass of transcription complexes and R-loops. *Nucleic Acids Res* 2020;**48**:10353–67.
25. Conti BA, Smogorzewska A. Mechanisms of direct replication restart at stressed replisomes. *DNA Repair (Amst)* 2020;**95**:102947. https://doi.org/10.1016/j.dnarep.2020.102947
26. Hobson DJ, Wei W, Steinmetz LM *et al.* RNA polymerase II collision interrupts convergent transcription. *Mol Cell* 2012;**48**:365–74. https://doi.org/10.1016/j.molcel.2012.08.027
27. Wilson MD, Harreman M, Taschner M *et al.* Proteasome-mediated processing of Def1, a critical step in the cellular response to transcription stress. *Cell* 2013;**154**:983–95. https://doi.org/10.1016/j.cell.2013.07.028
28. Fenstermaker TK, Petruk S, Kovermann SK *et al.* RNA polymerase II associates with active genes during DNA replication. *426 | Nat |* 2023;**620**:426–433. https://doi.org/10.1038/s41586-023-06341-9.
29. Bruno F, Coronel-Guisado C, González-Aguilera C. Collisions of RNA polymerases behind the replication fork promote alternative RNA splicing in newly replicated chromatin. *Mol Cell* 2024;**84**:221–33. https://doi.org/10.1016/j.molcel.2023.11.036
30. García-Pichardo D, Cañas JC, García-Rubio ML *et al.* Histone mutants separate R loop formation from genome instability induction. *Mol Cell* 2017;**66**:597–609. https://doi.org/10.1016/j.molcel.2017.05.014
31. Bayona-Feliu A, Herrera-Moyano E, Badra-Fajardo N *et al.* The chromatin network helps prevent cancer-associated mutagenesis at transcription-replication conflicts. *Nat Commun* 2023;**14**:6890.
32. Chong SY, Cutler S, Lin JJ *et al.* H3K4 methylation at active genes mitigates transcription-replication conflicts during replication stress. *Nat Commun* 2020;**11**:809. https://doi.org/10.1038/s41467-020-14595-4
33. Hao S, Wang Y, Zhao Y *et al.* Dynamic switching of crotonylation to ubiquitination of H2A at lysine 119 attenuates transcription–replication conflicts caused by replication stress. *Nucleic Acids Res* 2022;**50**:9873–92. https://doi.org/10.1093/nar/gkac734
34. Bayona-Feliu A, Barroso S, Muñoz S *et al.* The SWI/SNF chromatin remodeling complex helps resolve R-loop-mediated transcription–replication conflicts. *Nat Genet 2021 537* 2021;**53**:1050–63.
35. Bhowmick R, Mehta KPM, Lerdrup M *et al.* Integrator facilitates RNAPII removal to prevent transcription-replication collisions and genome instability. *Mol Cell* 2023;**83**:2357–66. https://doi.org/10.1016/j.molcel.2023.05.015
36. Papadopoulos D, Solvie D, Baluapuri A *et al.* MYCN recruits the nuclear exosome complex to RNA polymerase II to prevent transcription-replication conflicts. *Mol Cell* 2022;**82**:159–76. https://doi.org/10.1016/j.molcel.2021.11.002
37. St Germain CP, Zhao H, Sinha V *et al.* Genomic patterns of transcription–replication interactions in mouse primary B cells. *Nucleic Acids Res* 2022;**50**:2051–73. https://doi.org/10.1093/nar/gkac035
38. Tsirkas I, Dovrat D, Thangaraj M *et al.* Transcription-replication coordination revealed in single live cells. *Nucleic Acids Res* 2022;**50**:2143–56. https://doi.org/10.1093/nar/gkac069
39. Ginno PA, Lott PL, Christensen HC *et al.* R-Loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell* 2012;**45**:814–25. https://doi.org/10.1016/j.molcel.2012.01.017
40. Mátés L, Chuah MKL, Belay E *et al.* Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nat Genet* 2009;**41**:753–61.https://doi.org/10.1038/ng.343
41. Maldonado R, Schwartz U, Silberhorn E *et al.* Nucleosomes stabilize ssRNA-dsDNA triple helices in human cells. *Mol Cell* 2019;**73**:1243–54. https://doi.org/10.1016/j.molcel.2019.01.007
42. Wang J, Rojas P, Mao J *et al.* Persistence of RNA transcription during DNA replication delays duplication of transcription start sites until G2/M. *Cell Rep* 2021;**34**:108759. https://doi.org/10.1016/j.celrep.2021.108759
43. Schindelin J, Arganda-Carreras I, Frise E *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012;**9**:676–82. https://doi.org/10.1038/nmeth.2019
44. Lalonde M, Ummethum H, Trauner M *et al.* An automated image analysis pipeline to quantify the coordination and overlap of transcription and replication activity in mammalian genomes. *Methods Cell Biol* 2023;**182**:199–219. https://doi.org/10.1016/BS.MCB.2023.05.012
45. Sun Y, Duthaler S, Nelson BJ. Autofocusing in computer microscopy: selecting the optimal focus algorithm. *Microsc Res Tech* 2004;**65**:139–49. https://doi.org/10.1002/jemt.20118
46. Weigert M, Schmidt U, Haase R *et al.* Star-convex polyhedra for 3D object detection and segmentation in microscopy. In: *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*.Snowmass, CO, USA, 2020, 3655–62. https://doi.org/10.1109/WACV45572.2020.9093435
47. Schmidt U, Weigert M, Broaddus C *et al.* Cell detection with star-convex polygons. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2018;**11071 LNCS**:265–73.
48. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009;**25**:1754–60. https://doi.org/10.1093/bioinformatics/btp324
49. Danecek P, Bonfield JK, Liddle J *et alt al.* Twelve years of SAMtools and BCFtools. *Gigascience* 2021;**10**:giab008.https://doi.org/10.1093/gigascience/giab008
50. Lindstrand A, Eisfeldt J, Vezzi F *et al.* TIDDIT, an efficient and comprehensive structural variant caller for massive parallel sequencing data. *F1000Res* 2017;**6**:664.
51. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;**26**:841–2. https://doi.org/10.1093/bioinformatics/btq033
52. Lawrence M, Huber W, Pagès H *et al.* Software for computing and annotating genomic ranges. *PLoS Comput Biol* 2013;**9**:e1003118. https://doi.org/10.1371/journal.pcbi.1003118
53. Krueger F, James F, Ewels P *et al.t al.* FelixKrueger/TrimGalore: v0.6.10 - add default decompression path. *Zenodo* 2023. https://doi.org/10.5281/ZENODO.7598955
54. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;**9**:357–9. https://doi.org/10.1038/nmeth.1923
55. Broad Institute 2019; Picard Toolkit.
56. Ramírez F, Ryan DP, Grüning B *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 2016;**44**:W160–5. https://doi.org/10.1093/nar/gkw257

57. Amemiya HM, Kundaje A, Boyle AP. The ENCODE blacklist: identification of problematic regions of the genome. *Sci Rep* 2019;**9**:9354.
58. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;**15**:550. https://doi.org/10.1186/s13059-014-0550-8
59. Leek JT, Johnson WE, Parker HS *et al.* The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 2012;**28**:882–3. https://doi.org/10.1093/bioinformatics/bts034
60. Crossley MP, Bocek MJ, Hamperl S *et al.* qDRIP: a method to quantitatively assess RNA–DNA hybrid formation genome-wide. *Nucleic Acids Res* 2020;**48**:e84. https://doi.org/10.1093/nar/gkaa500
61. Dunham I, Kundaje A, Aldred SF *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;**489**:57–74.
62. Reverón-Gómez N, González-Aguilera C, Stewart-Morgan KR *et al.* Accurate recycling of parental histones reproduces the histone modification landscape during DNA replication. *Mol Cell* 2018;**72**:239–49. https://doi.org/10.1016/j.molcel.2018.08.010
63. Zhang Y, Liu T, Meyer CA *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008;**9**:R137. https://doi.org/10.1186/gb-2008-9-9-r137
64. Liu JH, Xi K, Zhang X *et al.* Structural flexibility of DNA-RNA hybrid duplex: stretching and twist-stretch coupling. *Biophys J* 2019;**117**:74–86. https://doi.org/10.1016/j.bpj.2019.05.018
65. Kowarz E, Löscher D, Marschalek R. Optimized sleeping beauty transposons rapidly generate stable transgenic cell lines. *Biotechnol J* 2015;**10**:647–53. https://doi.org/10.1002/biot.201400821
66. Zhou J, Snyder AR, Lieberman PM. Epstein-Barr virus episome stability is coupled to a delay in replication timing. *J Virol* 2009;**83**:2154–62. https://doi.org/10.1128/JVI.02115-08
67. Carroll SM, Trotter & Geoffrey J, Wahl MM. Replication timing control can be maintained in extrachromosomally amplified genes. *Mol Cell Biol* 1991;**11**:4779–85.
68. Lam FC, Kong YW, Huang Q *et al.* BRD4 prevents the accumulation of R-loops and protects against transcription–replication collision events and DNA damage. *Nat Commun* 2020;**11**:4083.
69. Helmrich A, Ballarino M, Tora L. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol Cell* 2011;**44**:966–77. https://doi.org/10.1016/j.molcel.2011.10.013
70. Okamoto Y, Abe M, Itaya A *et al.* FANCD2 protects genome stability by recruiting RNA processing enzymes to resolve R-loops during mild replication stress. *FEBS J* 2019;**286**:139–50. https://doi.org/10.1111/febs.14700
71. Shao X, Joergensen AM, Howlett NG *et al.* A distinct role for recombination repair factors in an early cellular response to transcription–replication conflicts. *Nucleic Acids Res* 2020;**48**:5467–84. https://doi.org/10.1093/nar/gkaa268
72. Said M, Barra V, Balzano E *et al.* FANCD2 promotes mitotic rescue from transcription-mediated replication stress in SETX-deficient cancer cells. *Commun Biol* 2022;**5**:1395. https://doi.org/10.1038/s42003-022-04360-2
73. Okamoto Y, Iwasaki WM, Kugou K *et al.* Replication stress induces accumulation of FANCD2 at central region of large fragile genes. *Nucleic Acids Res* 2018;**46**:2932–44. https://doi.org/10.1093/nar/gky058
74. Groelly FJ, Dagg RA, Petropoulos M *et al.* Mitotic DNA synthesis is caused by transcription-replication conflicts in BRCA2-deficient cells. *Mol Cell* 2022;**82**:3382–97. https://doi.org/10.1016/j.molcel.2022.07.011
75. Castellano-Pozo M, Santos-Pereira J, Rondón AG *et al.* R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Mol Cell* 2013;**52**:583–90. https://doi.org/10.1016/j.molcel.2013.10.006
76. Huynh MT, Sengupta B, Krajewski WA *et al.* Effects of Histone H2B Ubiquitylations and H3K79me3 on Transcription Elongation. *ACS Chem Biol* 2023;**18**:537–48. https://doi.org/10.1021/acschembio.2c00887
77. Wood K, Tellier M, Murphy S. DOT1L and H3K79 methylation in transcription and genomic stability. *Biomol* 2018;**8**:11.
78. Kari V, Raul SK, Henck JM *et al.* The histone methyltransferase DOT1L is required for proper DNA damage response, DNA repair, and modulates chemotherapy responsiveness. *Clin Epigenet* 2019;**11**:4. https://doi.org/10.1186/s13148-018-0601-1
79. Kang JY, Park JW, Hahm JY *et al.* Histone H3K79 demethylation by KDM2B facilitates proper DNA replication through PCNA dissociation from chromatin. *Cell Prolif* 2020;**53**:e12920. https://doi.org/10.1111/cpr.12920
80. Jones B, Su H, Bhat A *et al.* The Histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLos Genet* 2008;**4**:e1000190. https://doi.org/10.1371/journal.pgen.1000190
81. Daigle SR, Olhava EJ, Therkelsen CA *et al.* Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. *Blood* 2013;**122**:1017–25. https://doi.org/10.1182/blood-2013-04-497644
82. Wu A, Zhi J, Tian T *et al.* DOT1L complex regulates transcriptional initiation in human erythroleukemic cells. *Proc Natl Acad Sci USA* 2021;**118**:e2106148118. https://doi.org/10.1073/pnas.2106148118
83. Dunn K, Griffith JD. The presence of RNA in a double helix inhibits its interaction with histone protein. *Nucl Acids Res* 1980;**8**:555–66. https://doi.org/10.1093/nar/8.3.555
84. Castillo-Guzman D, Chédin F. Defining R-loop classes and their contributions to genome instability. *DNA Repair (Amst)* 2021;**106**:103182. https://doi.org/10.1016/j.dnarep.2021.103182
85. Enright HU, Miller WJ, Hebbel RP. Nucleosomal histone protein protects DNA from iron-mediated damage. *Nucl Acids Res* 1992;**20**:3341–6. https://doi.org/10.1093/nar/20.13.3341
86. Brambilla F, Garcia-Manteiga JM, Monteleone E *et al.* Nucleosomes effectively shield DNA from radiation damage in living cells. *Nucleic Acids Res* 2020;**48**:8993–9006. https://doi.org/10.1093/nar/gkaa613
87. García-Rubio M, Aguilera P, Lafuente-Barquero J *et al.* Yra1-bound RNA–DNA hybrids cause orientation-independent transcription–replication collisions and telomere instability. *Genes Dev* 2018;**32**:965–77. https://doi.org/10.1101/gad.311274.117
88. Sankar TS, Wastuwidyaningtyas BD, Dong Y *et al.* The nature of mutations induced by replication–transcription collisions. *Nature* 2016;**535**:178–81.
89. Macheret M, Halazonetis TD. Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress. *Nature* 2018;**555**:112–6. https://doi.org/10.1038/nature25507
90. Kotsantis P, Silva LM, Irmscher S *et al.* Increased global transcription activity as a mechanism of replication stress in cancer. *Nat Commun* 2016;**7**:13087.
91. Tous C, Aguilera A. Impairment of transcription elongation by R-loops in vitro. *Biochem Biophys Res Commun* 2007;**360**:428–32. https://doi.org/10.1016/j.bbrc.2007.06.098
92. Huyen Y, Zgheib O, DiTullio RA *et al.* Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nat 2004 4327015* 2004;**432**:406–11.
93. Salguero C, Valladolid C, Robinson HMR *et al.* Targeting ATR in cancer medicine. *Cancer Treat Res* 2023;**186**:239–83. https://doi.org/10.1007/978-3-031-30065-3_14
94. Carrasco-Salas Y, Malapert A, Sulthana S *et al.* The extruded non-template strand determines the architecture of R-loops. *Nucleic Acids Res* 2019;**47**:6783–95. https://doi.org/10.1093/nar/gkz341

Received: July 3, 2024. **Revised:** January 31, 2025. **Editorial Decision:** February 3, 2025. **Accepted:** February 9, 2025